Author:
Jen, Min-Hua

Title:
Health outcomes and income inequality : a multilevel analysis of the Wilkinson hypothesis

Health outcomes and income inequality: a multilevel analysis of the Wilkinson hypothesis

Min-Hua Jen

A dissertation submitted to the University of Bristol in accordance with the requirements of the Degree of Doctor of Philosophy in the Faculty of Social Science

School of Geographical Sciences

University of Bristol

September 2006

Abstract

This thesis aims to evaluate Richard Wilkinson's arguments that, in developed economies, it is not (absolute) income that is a main determining factor of health but the degree of inequality of income within a society. The distinctive nature of this thesis is twofold. This is the first study to analyse individual data at the scale of countries which was the scale used originally used by Wilkinson. Moreover, random coefficient modelling is used to model individual and country variables simultaneously and thereby overcome the ecological fallacy that has troubled previous research.

The study is based on four distinct pieces of research, which are international-based comparisons of aggregated and individual level data through time. One is a longitudinal analysis of life expectancy at birth for some 196 countries across thirty years. Group trajectory modelling is used to reveal groups of countries with distinctive trends and these trends are then related to changing GDP and income inequality. Using this aggregated data, support is found for the Wilkinson hypothesis in that in developed countries life-expectancy is related to GDP but not to income inequality. The other three studies are based on individual level data derived from the World Values Survey. These analyses provide a direct test of the Wilkinson hypothesis using micro data on individuals, and macro data on relative inequalities analysed simultaneously. They investigate the individual self-rated health associated with country income and inequality data, while also estimating the relationship between subjective well-being (happiness and life satisfaction) and individual income and country inequality. Finally, they uncover the underling relationship between individual self-rated health and individual and country level social trust as an evaluation of the hypothesis that social cohesion affects human health. In summary, with this improved methodology it is found that the Wilkinson hypothesis is not supported in terms of income inequality, but there is an effect for social trust, over and above individual factors.

Dedication and Acknowledgement

I would like to take this opportunity to thank several people who were influential in my research or helped in other ways. First I should say Kelvyn Jones, Richard Harris, Leslie Hepple and Ron Johnston for guiding me, answering my many questions, listening to my ideas. I would also like to thank Jon Rasbash and Edward Thomas for sorting out my technical problems, and Tony Hoare and Jenny Arnott for helping to proof read this thesis and correcting my idiosyncratic English. Any mistakes that remain are purely my own fault. There are of course many other people whom I have met during my PhD, to numerous to mention, but who have all contributed in their own way and to whom I am eternally grateful.

I also owe a debit of gratitude to the Chiang Ching-Kuo Foundation for providing me with a scholarship to finish my PhD, and to my family who encouraged me to pursue my dreams and who were always there when I needed them. In particular, I would like to dedicate this thesis to my father who gave me a lot of support mentally and financially.

感謝我親愛的家人一路以來的默默的支持
尤其是我的爸媽 總是在電話那頭傾聽
分享我的歡樂 撫慰我的憂愁
縱使千言萬語 也無法完全傳達我的感激
希望將此論文獻給我親愛的爸媽

And a special thank you to my lovely boyfriend Davy who is always there for me no matter good time or hard time. There is no word to express my grateful and how lucky I am to have him in my life.

I declare that the work in this dissertation was carried out in accordance with the Regulations of the University of Bristol. The work is original, except where indicated by special reference in the text, and no part of the dissertation has been submitted for any other academic award. Any views expressed in the dissertation are those of the author.

SIGNED: .... DATE:...................................

Contents

List of Figures

List of Tables

Chapter 1
Introduction

The last ten years have seen a major interest in the macro determinants of health. Although this interest is not a new one (for example, Ancel Keys planned his ground-breaking the Seven Countries Study which was the first to examine systematically the relation among lifestyle, diet, and the rates of heart attack and stroke in contrasting population in the late 1950s) it recently has received increased attention as the effectiveness of an individual 'lifestyle' approach to health has been questioned (Shy, 1997).

A major proponent of the idea that it is not the characteristics of the individual but of the society in which they live that determine ill health has been raised by Richard Wilkinson, particularly in his 1996 book *Unhealthy Societies* and in the 2001 follow-up *Mind the Gap*. The detail of his arguments is considered in Chapter Two of this thesis; in short, Wilkinson argues that the major cause of ill-health in advanced economies is not lifestyle, environmental or genetic factors. Instead, the causes are societal. Put simply, in the rich countries of the West, unequal societies produce injurious levels of social stress that result in premature death. Health inequalities therefore are seen as a result of relative income inequalities undermining social cohesion and exacerbating psychosocial processes of stress. The central focus of this thesis is to evaluate Wilkinson's hypothesis.

Although there are an enormous number of papers evaluating and challenging Wilkinson's arguments, none of them assesses the underlying relationship between income inequality and health outcomes across a wide range of countries, using data at the individual as well as the aggregate level. Consequently, they either conflate or ignore micro and macro, as well as geographical perspectives on health. The distinctive nature of this thesis is threefold:

1) An analysis is undertaken of the relationship between health and income inequality using the same macro-scale of analysis as used by Wilkinson (1996, 2001). Whilst other researchers have explored the Wilkinson hypothesis at the

sub-national scales of States and Metropolitan Areas within the USA, here extensive country level data are used – the same definition of 'societies' used by Wilkinson when providing empirical evidence for his arguments.

2) An appropriate multilevel methodology is developed to model simultaneously at the macro-, country level as well as at the micro-, individual level, permitting the relative importance of micro and macro effects on health to be quantified and the spurious effects of data aggregation to be avoided.

3) The study has a longitudinal perspective, modelling changes over time.

The structure of the thesis is as follows. Chapter Two looks in detail at the underlying arguments of and evidence for the Wilkinson thesis. In particular, it reviews previous studies of the 'relative income hypothesis' – the idea that it is not individual (absolute) income which affects people's health but income relative to others within a society. The chapter also reviews criticisms of this income inequality hypothesis. In particular, it focuses on Gravelle's (1998) argument that the relationships found by Wilkinson (at the country level) are an artefact of a non-linear relationship at the individual level between income and health, which, once aggregated, cause a statistically significant but spurious association between income inequality and people's average health to emerge.

In the chapter, a simulation confirms the problem and demonstrates the need for a multilevel strategy to re-examine Wilkinson's hypothesis. Summarised are recent findings based on the multilevel modelling of associations between health, income and income inequality at a variety of scales such as communities, regions or states; and of associations between health and the mediating factor of social cohesion. The chapter concludes with an outline of the research questions to be pursued in the following chapters to validate or to challenge Wilkinson's hypothesis, and also considers the nature of the data required to answer those questions.

The rest of the thesis is based on four distinct but complementary pieces of research, which are international-based comparisons of health outcomes through time. Chapter Three presents a longitudinal analysis of some 196 countries' changing life expectancies. This modelling is based on the collation of a dataset for 1970-2002, that has information on life expectancy, as well as measures of income and income inequality. The analysis is used to uncover the basic trends in mortality between countries, using comparable data, and to relate these trends to changing levels of income and income inequality.

There are three aims for this part of the study: firstly to evaluate the general patterns through time and in a wide range of different countries to assess the degree of between occasion (year) and between country variation. Secondly, to identify subgroups of countries which have distinctive developmental patterns. Finally, examine the changing patterns of mortality as a function of a country's classification of income and income inequality. In methodological terms, the recently developed semi-parametric random coefficient approach of Nagin (2005) is used to identify countries with similar trajectories. This methodology will also be used to predict what would have happened to mortality trends under a range of circumstances, if countries had achieved the same lack of income inequality enjoyed by Sweden. The results of this study previously have been presented at the Methodology of Longitudinal Survey conference (University of Essex, 2006) and at the Emerging and New Research in Geographies of Health and Impairment (University of Maynooth, Ireland, 2004).

Chapters Four, Five and Six are all based on individual level data obtained from the World Values Survey (Inglehart, 1997). Individual data are vital if we are to avoid the ecological fallacy associated with aggregate data analysis – which is, as will be shown, a key flaw inherent to Wilkinson's empirical work. The underlying aims of these three studies is to use data collected on a comparable basis and using scientific sampling strategies in a large number of countries to provide a direct test of the Wilkinson hypothesis using micro data on individuals, and macro data on relative income inequalities at the country scale.

More specifically, Chapter Four reports the results of a multilevel analysis relating self-rated health at the individual level to individual income, and simultaneously to average income and income inequality at the country level. Given that Wilkinson developed his hypothesis in relation to countries as societies, so it is crucial to examine a wide variety of countries with a wide range of average income and income inequalities.

Wilkinson argues that once countries reach a certain stage of development (which he defines as GNP $5000 per capita in 1990) inequality takes over from per capita income as the primary determinant of health. Therefore it is also important that: (a) his hypothesis is examined over time, as countries become more developed; and (b) that the effect of income inequality on health within the model be allowed to vary between non- and advanced economies. Methodologically, this chapter considers the data provided by the successive waves of the World Values Survey and how this can be handled by a multilevel multinomial logit model.

In his book '*Unhealthy societies*' (1996), Wilkinson concluded that countries where there is greater inequality, experience higher social stress, greater unhappiness, poorer life satisfaction and consequently poorer health. In chapter 5, a set of models is developed to further examine the Wilkinson hypothesis, relating individual and country income variables to a set of outcomes representing self-reported health, happiness and life satisfaction. This allows an assessment of the effects of income inequality on a range of life indicators and also provides an assessment of the correlation between these different outcomes both between countries and within countries (where they are between people). Methodologically, this chapter uses a multivariate multilevel model. Preliminary results from this chapter were presented at the conference of ESRC Research Method in the University of Durham 2004.

As has been stated, Wilkinson contends that the key mechanisms between income distribution and health are psycho-social processes. He stresses that "what matters most ...probably even for people who still suffer from air pollution and damp housing is psycho-social welfare" (2000, 7). To him, the primary cause of health inequalities is not material disadvantage but psycho-social stress arising from perceived or actual relative disadvantage. Income inequality is said to undermine social cohesion which has a detrimental psychosocial effect on individual health (Wilkinson, 1996, 1998). Chapter 6 is a direct test of this hypothesis, providing the results of a set of models relating individual self-rated health to social trust, on individual and country levels, and taking account of individual social demographic and economic variables by multilevel modelling. The final modelled result is compared with Subramanian's *et al.*'s (2004) research which was undertaken within US communities in different scale rather than between countries.

Chapter Seven concludes the thesis. By summarising the findings of the preceding chapters and by drawing together the empirical evidence that has been obtained, a final assessment is made upon Wilkinson's hypothesis concerning the association between individual health outcomes and the absolute and relative income of a society. Comment is also made on the limitations of the research and on how the work may be developed in the future.

Chapter 2

A critical review of Wilkinson's psychosocial perspective on health inequalities

2.1 INTRODUCTION

There is a considerable debate about the importance of income and inequality in determining health outcomes. In particular, it has been argued by Richard Wilkinson in a series of papers and books, that in developed economies, it is not (absolute) income that is a main determining factor of health but the degree of inequality of income within a society. Furthermore, he stresses that the key mechanism causing health inequalities is not material disadvantage but psycho-social stress arising from perceived relative disadvantage.

This chapter reviews previous work on the Wilkinson's hypothesis and subjects the hypothesis to an initial critique. This review will spell out the argument largely in Wilkinson's own words and diagrams. The chapter has the following structure. Firstly, there is a detailed review of the income-inequality hypothesis that is put forward by Wilkinson, namely that the extent of dispersion of a place's income plays a more important role in determining health outcome than real income, especially in developed countries. Secondly, the chapter considers recent critiques of the hypothesis and examines the inconsistency and shortcomings of both methodology and data.

Thirdly, the key features of an appropriate methodology required to evaluate Wilkinson's hypothesis are examined: in particular, the advantages of using a multilevel modelling approach to assess the data and discover the true underlying relationship are considered. In this context attention is focused on Gravelle's (1998) argument that the inequality effect is an artefact of aggregating data from a non-linear individual-level relation between income and health. A simulation will be carried out to show the importance of Gravelle's argument and therefore the compelling need for a multilevel modelling when evaluating the Wilkinson hypothesis. The more recent multilevel studies that have been undertaken are then reviewed before finally listing the purpose and research questions for each subsequent chapter for evaluating Wilkinson's hypothesis.

2.2 PREVIOUS RESEARCH

2.2.1 Income inequality debates; Wilkinson's argument

The key to understanding Wilkinson's argument is to distinguish between absolute and relative income: the former relates to how much income an individual receives; the latter is the degree of inequality in a society or a place. A number of studies of developed economies have indicated that the higher an individual's income the better their health; that is the 'absolute income hypothesis' (Adler, 1993; Pritchett *et al*, 1996 and Ben-Shlomo *et al.*, 1996). However, Wilkinson (1996) in the book '*Unhealthy societies*' (US), stresses the relative income hypothesis, concluding that

"[t]he scale of income differences in a society is one of the most powerful determinants of health standards in different countries and that it influences health through its impact on social cohesion" (US, 1996, pix);

and that

*"[i]ncome **distribution** is linked to social cohesion which in turn is linked to mortality"* (US, 1996, ix, emphasis added)

Moreover, Wilkinson (2000) in his later book '*Mind the gap*' (MTG), challenges the materialist explanation that individual income simply buys access to certain goods:

"[m]ost of us working in this area assumed that they [i.e. health inequalities] *must be due to differences in material living standards between social classes. It looked then as if our task was to discover the damage done by bad housing, air pollution, poor diet, inadequate housing and other aspects of people's circumstances... we assumed the key relationships were those between people and things"* (MTG, 2000, p6-7).

He continues:

*"[w]hat matters most...probably even for people who still suffer from air pollution and damp housing is **psycho-social welfare**"* (MTG, 2000, p7, emphasis added).

There are two important points which have to be emphasized here. Firstly, relative income differences are the effect of places (societal/geographical), which are above and beyond individual income effects on health. Secondly, for Wilkinson, it is not materialist but psycho-social factors that make people ill. These are the key elements that constitute the 'relative income hypothesis'.

Wilkinson bases this relative income hypothesis on a large and complex empirical evidence base, now reviewed in the following order: the limited role of individual risk factors in explaining health inequality (evidence 1); the contrasting relationship for income and health found between-countries and within-countries (evidence 2 and 3); and then specific examples of places with high social cohesion having good experiences of health and other social outcomes (evidence 4 and 5). This empirical evidence will then allow me to spell out the effects of income inequality on health via social cohesion as the psychosocial pathway in an overall model (evidence 6 and 7).

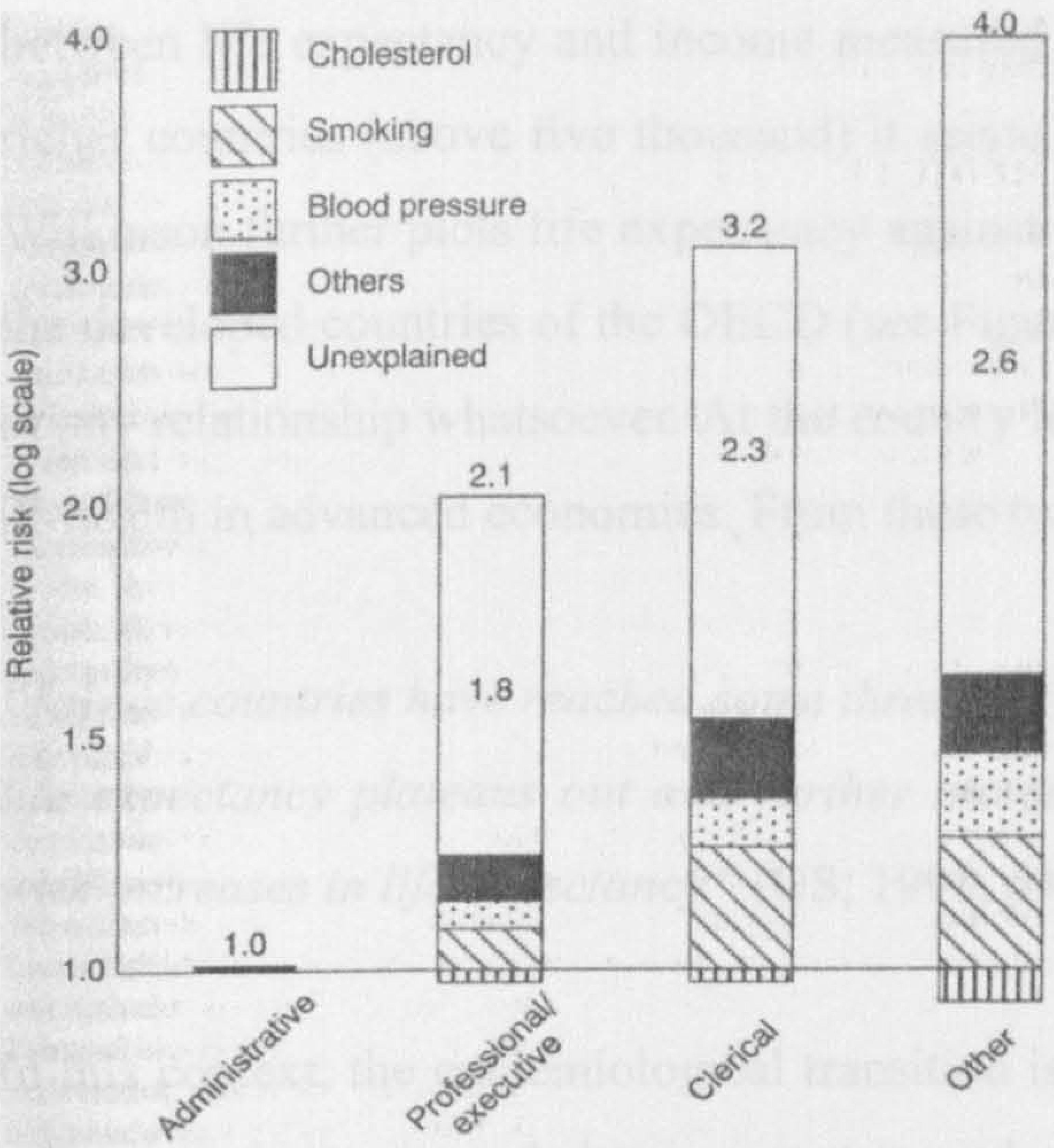
2.2.1.1 Evidence 1: The low-role of individual risk factors

In the long history of studying health inequality, it has been suggested that the causes of health differences are to be found in differences in smoking, drinking, diet, exercise and the like. Several large studies have measured the contribution of individual behavioural factors such as these to differentials in death rates. In the British context, one of the major studies has been the Whitehall study (Marmot *et al.*, 1991) which is a prospective longitudinal study of British civil servants. Figure 2.1 summarises what this study has found. It shows the relative risk of mortality from coronary heart disease on each discrete bar chart representing different employment class or grade in the civil service. The health inequalities are large with a four-fold difference between "other" (that is the lowest grade) and the highest administrative grade. The graph also shows the extent of differences that can be statistically explained by various risk factors. Especially, only one third can be explained by individual risk factors even in combination, thereby suggesting that the majority of heath inequality variation must be explained by something other than individual behaviour.

The role of individual risk factors has also been brought into question in the work of Rose (1981). He calculated the impact of a low cholesterol diet on individual risk of heart attack using data from the Framingham study (a key and pioneering American-based study) and made the following statement:

"[i]f men were to modify their diet enough to reduce their cholesterol by 10 percent up to the age of 55, 98 per cent of them would have to eat differently every day for forty years without having prevented a heart attack by doing so. If the pay-offs are so small even for preventive measures, which are, regarded as most important what hope is there for less important risk factors in less important causes of death" (US, 1996, p71).

Figure 2.1 Relative risk of death from coronary heart disease according to employment grade, and proportions of differences that can be explained statistically by various risk factors



Source: Wilkinson, (US, 1996, p591)

Wilkinson also dismisses other individual-based explanations for health inequalities:

"[i]mportant differences in health...cannot be attribute to genetics, only a small part is played by healthier people moving up the social ladder and less healthy moving down; .. almost none of the health differences results from differences in medical care; the

differences are only partly explained by health-related behaviour....the main explanations of health inequalities have to be found in the effects of the different social and economic circumstances in which people live" (US, 1996, p7).

Consequently, it is not individual risks that relate to specific diseases. Instead

"[t]he most important single part of the picture ...concerns the social determinants of general vulnerability" (MTG, 2000, p9).

2.2.1.2 Evidence 2: GDP (income) and health: between-country relations

Wilkinson (in MTG, 2000) then continues his argument by looking at macro factors that could determine health inequality. In Figure 2.2, Wilkinson plots life expectancy against income per capita for a range of countries in four different eras. In countries with average income per person below five thousand dollars, there is a strong direct relationship between life expectancy and income measured as GNP PPP pc[1]. However, among those richer countries (above five thousand) it seems income has no effect on life expectancy. Wilkinson further plots life expectancy against income (represented by GNP PPP pc) for the developed countries of the OECD (see Figure 2.3) and concludes there is no evidence of any relationship whatsoever. At the country level, absolute income appears unrelated to ill-health in advanced economies. From these two figures he concludes that

"[o]nce countries have reached some threshold of income (around $5k per capita in 1990) life expectancy plateaus out and further increases in GNPpc are no longer associated with increases in life expectancy" (US, 1996, p34).

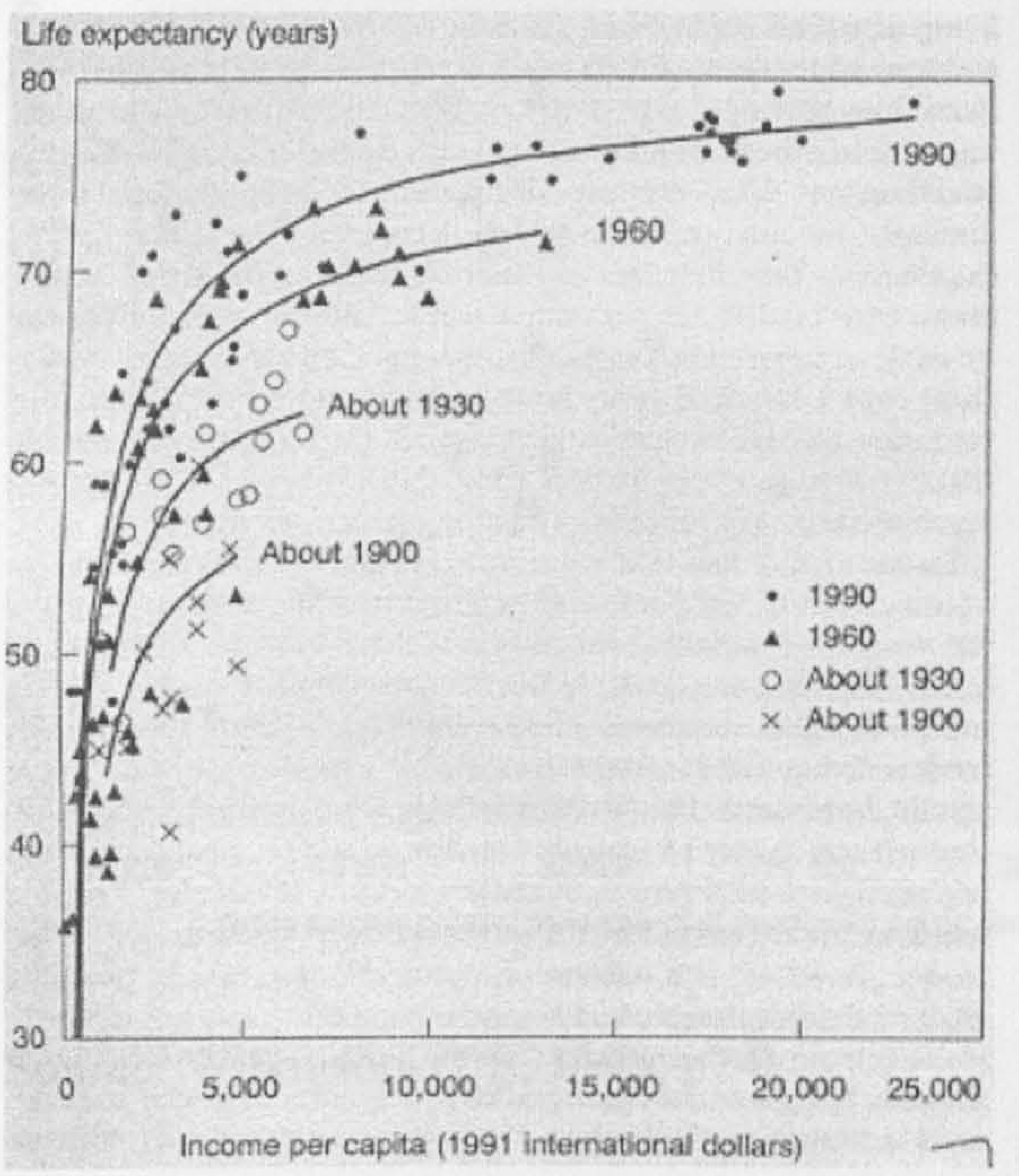
In this context, the epidemiological transition is seen in a new light, in that both the life expectancy experienced by a country and the transition from the dominance of communicable disease (e.g. tuberculosis) to degenerative disease (e.g. heart disease) is said to be determined by the income of a country. Once living standards are in excess of a minimum standard that is adequate to ensure basic material standards for all, there is an

1 GNP PPP pc: GNP pc is Gross National Product per capita. It represents the net final product of total goods and service at market value for an economy shared by its population, including the residents and foreign labour within the territory for a period of time. PPP is a rate of exchange that accounts for price differences across countries allowing international comparisons at the PPP US$ rate. A PPP US$1 has the same purchasing power in the domestic economy as $1 has in the United States. (United Nations, 2001)

epidemiological transition where the main cause of death shifts from infectious disease to degenerative disease. Wilkinson writes

"[r]icher countries... have advanced beyond a crucial stage when living standards reached a threshold level adequate to ensure basic material standards for all. This is marked by the epidemiological transition when infectious diseases gave way to cancer and degenerative diseases as the main causes of death" (US, 1996, p2).

Figure 2.2 Changing life expectancy and income per capita



Source: Wilkinson, (US, 1996, p34)

He also claims that

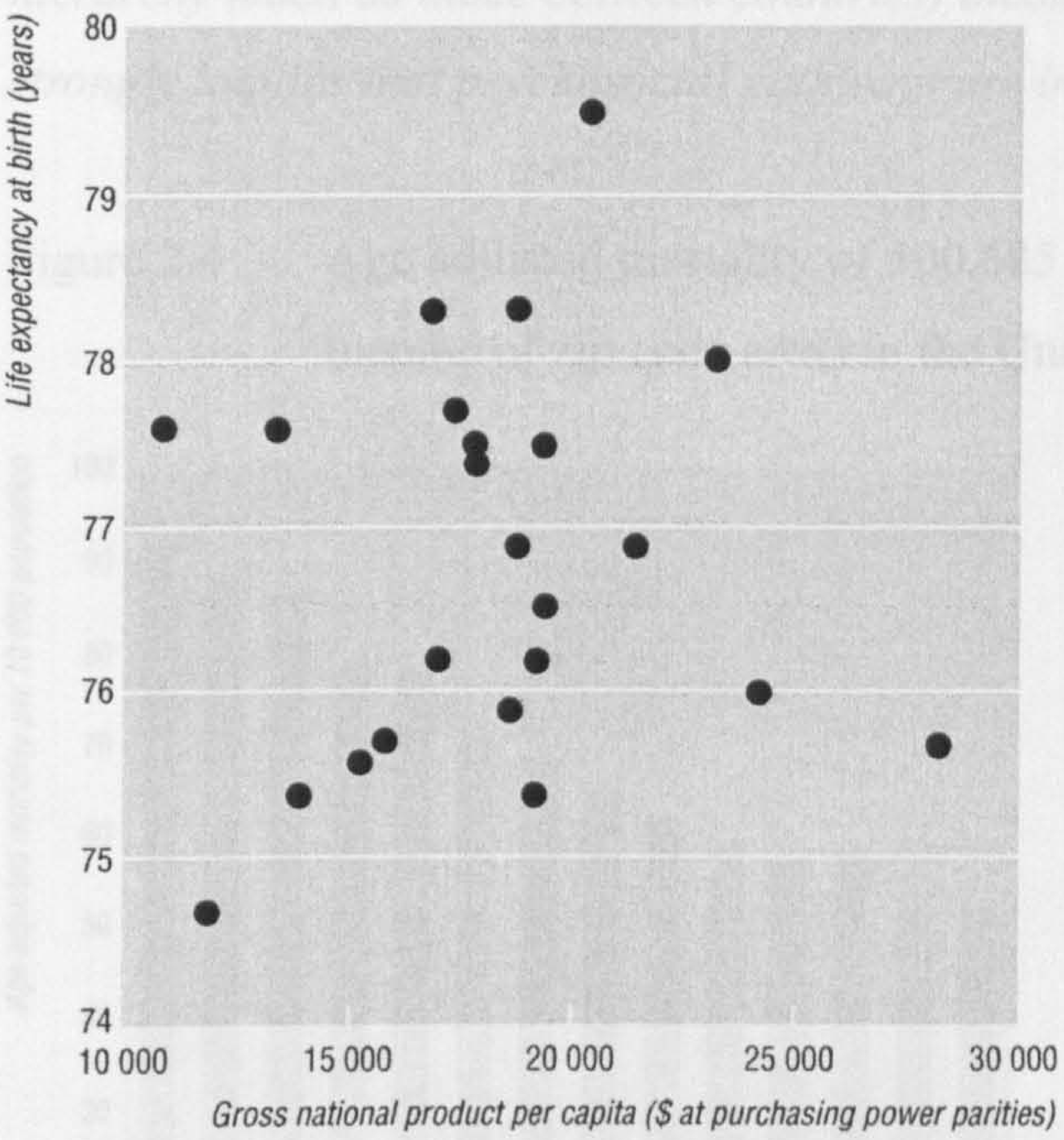
"[d]uring the epidemiological transition the so-called affluent diseases became the diseases of the poor in affluent societies" (US, 1996, p44);

and that

"[t]he transition from infectious to degenerative causes of death marks the attainment of the minimum standard consistent with health among the vast majority of the population" (US, 1996, p7).

So below a threshold of around $5K per capita in 1990, material disadvantage is a key process, communicable diseases are dominant, and life is relatively short. Above the threshold, material disadvantage appears not to be important.

Figure 2.3 Relationship of life expectancy and gross national product per capita in OECD countries, 1993



Source: Wilkinson, (US, 1996, p74)

2.2.1.3 Evidence 3: income and health: within-country relations

Wilkinson notes that within developed countries such as the USA, the annual age adjusted death rates are anywhere between twice and four times as high for poorer as richer people (see Figure 2.4):

"[t]he usual pattern is a continuous gradient across the whole society" (MTG, 2001, p5).

The central crux of his argument here is that income within a society has to be given a relative interpretation:

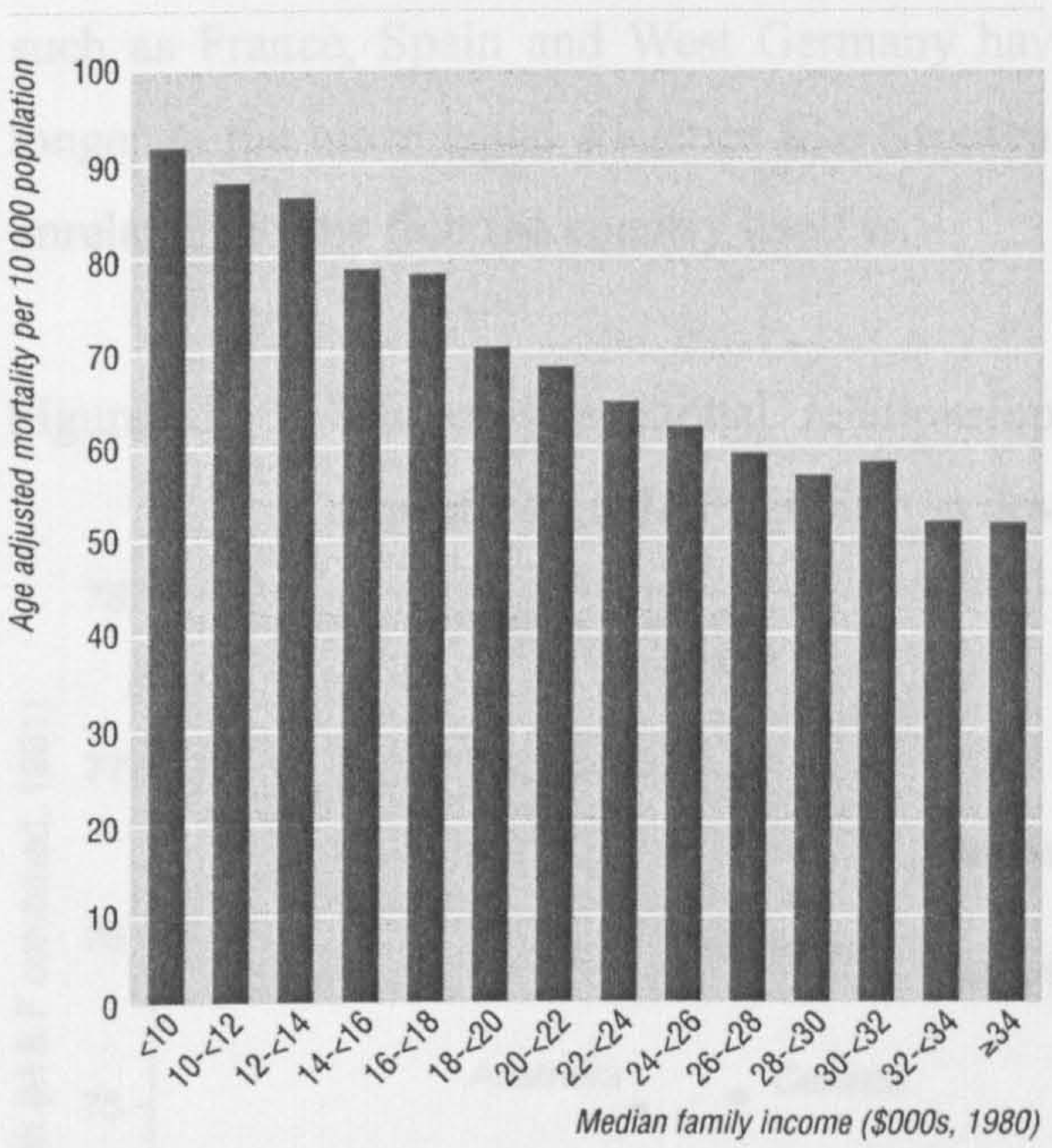
"[h]ealth is related to differences in living standards within developed societies but not to differences between them. This inconsistency is resolved by evidence suggesting that what

matters within countries is relative rather than absolute income levels" (US, 1996, p7).

Moreover,

"[w]here income is related to social status, as it is within countries, it is also related to health. Where income differences mean little or nothing for people's position in the social hierarchy (such as those between countries) income makes little difference to health. This strongly implies that psychosocial pathways are important" (MTG, 2000, p10-11).

Figure 2.4 Age adjusted mortality of 300,685 white American men by median family income of zip code areas in the United States



Source: Wilkinson, (US, 1996, p73)

In summary,

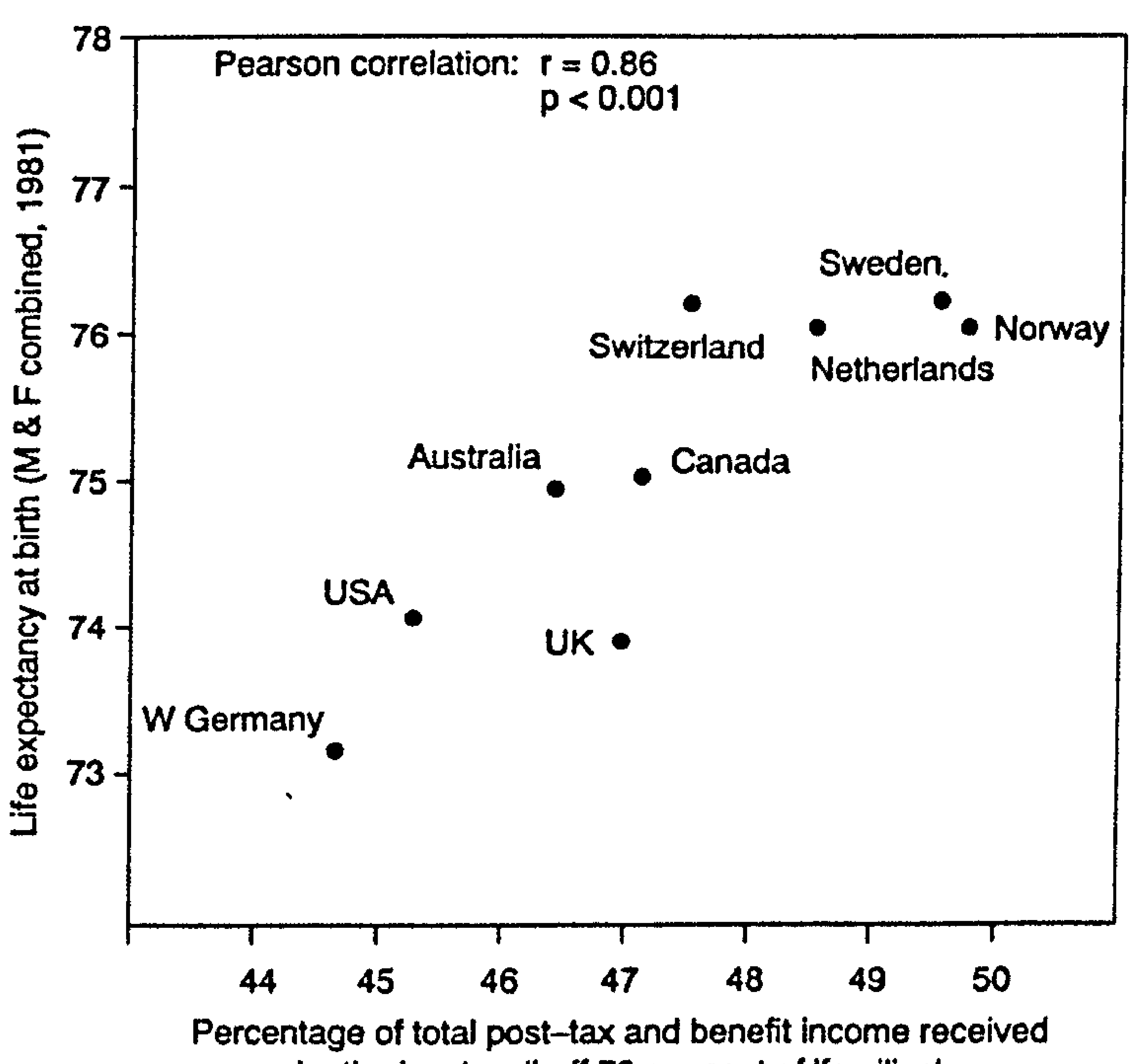
"[i]t now looks as if individual income like inequality affects health though psychosocial processes" (MTG, 2000, p12).

For the relative income argument, in developed countries, it is not what income you have that determine health outcomes, but income you have in relation to others in your society; not total income but comparative status.

2.2.1.4 Evidence 4: Income inequality and Health

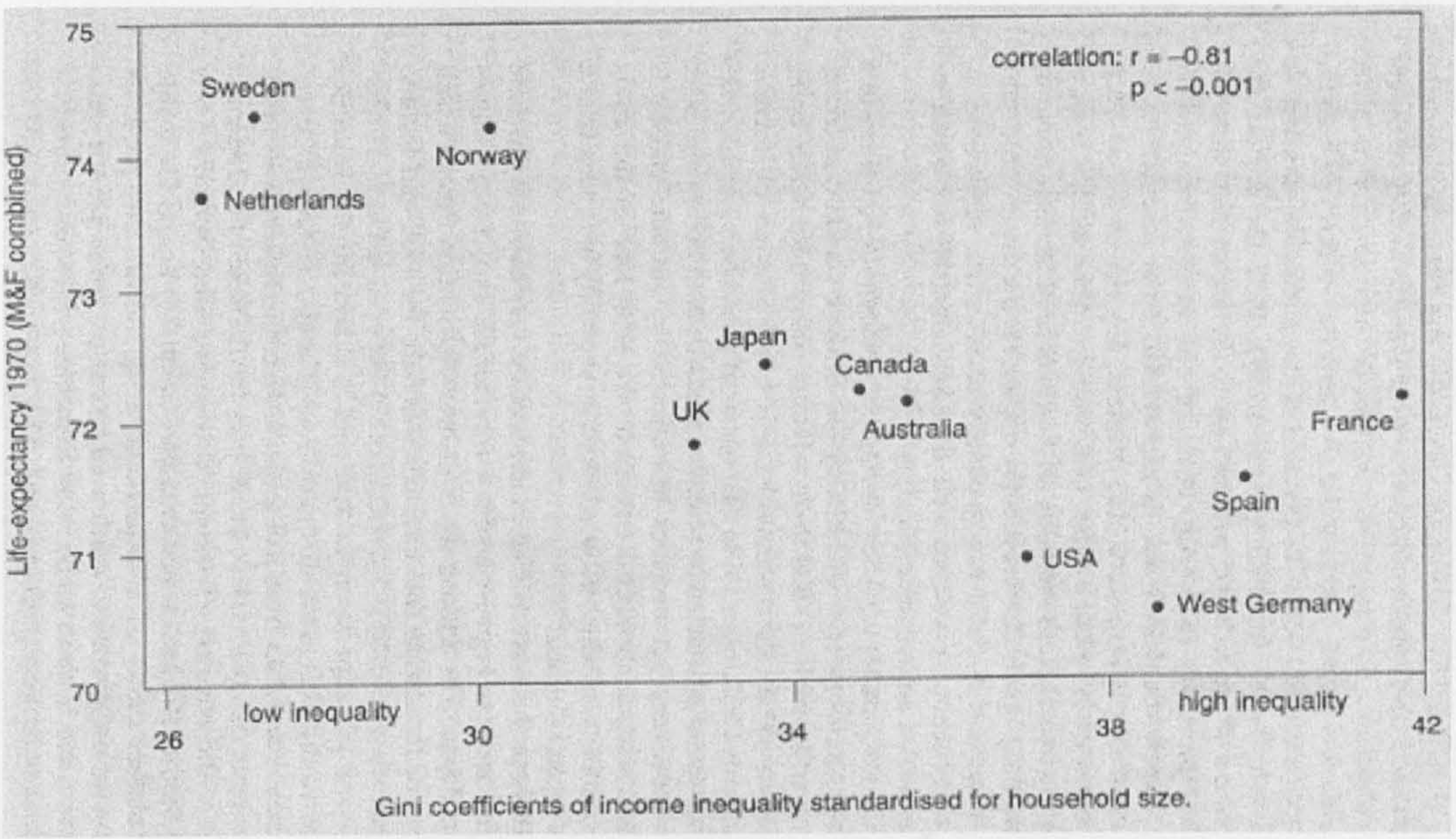
More direct evidence of the relative income hypothesis comes from examining the relationship between health outcomes and relative income within countries. In Figure 2.5, there is a positive and significant correlation of 0.86 ($P < 0.001$) between the share of total personal income received by the least well-off 70 per cent of households in OECD countries and life expectancy. In relatively egalitarian societies where the poorest 70% receive around 50% of total income (such as Sweden and Norway), life expectancy is around 76 years; thus compares favourably to the former West Germany and USA which are less egalitarian, and have a shorter life expectancy by some two to three years. Similarly, when Wilkinson plots life expectancy against the Gini measure of inequality (see Figure 2.6) he finds a clear relationship ($r = -0.81$): unequal, developed societies such as France, Spain and West Germany have the worse life experience; people live longer in the more equal societies like Sweden and the Netherlands. The relationship is unrelated to how rich the country itself is.

Figure 2.5 The cross-sectional relationship between income distribution and life expectancy (M&F) at birth in developed countries



Source: Wilkinson, (US, 1996, p76)

Figure 2.6 Inequality and life expectancy for developed countries



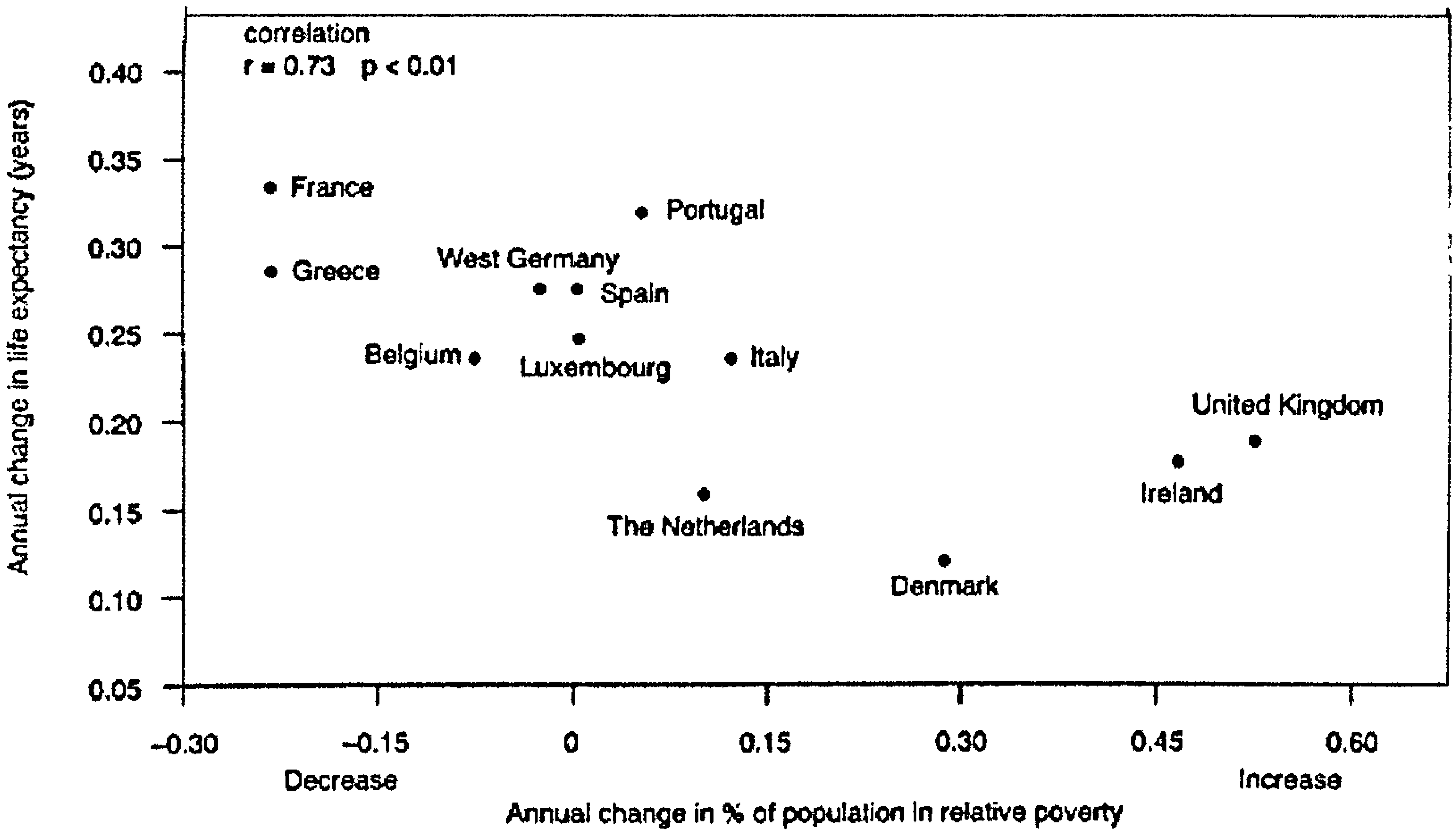
Source: Wilkinson, (US, 1996, p84)

Moreover, the relationship between the *changes* of life expectancy and population proportion in poverty between 1975 and 1985, of the twelve European countries in Figure 2.7 shows a significant and negative correlation of 0.73 ($P < 0.01$). Reducing the relative poverty of populations such as in France and Greece increased life expectancy through time.

The impact of income inequality on health not only appears internationally but also holds within nations such as the United States, as shown in Figure 2.8. The age adjusted mortality for USA states is highly correlated with the percentage share of household income received by the poorest 50 percent of the population (0.62; $P < 0.001$). It is still very significant even after taking into account of average incomes, absolute poverty, racial differences and smoking in an aggregate analysis of USA states.

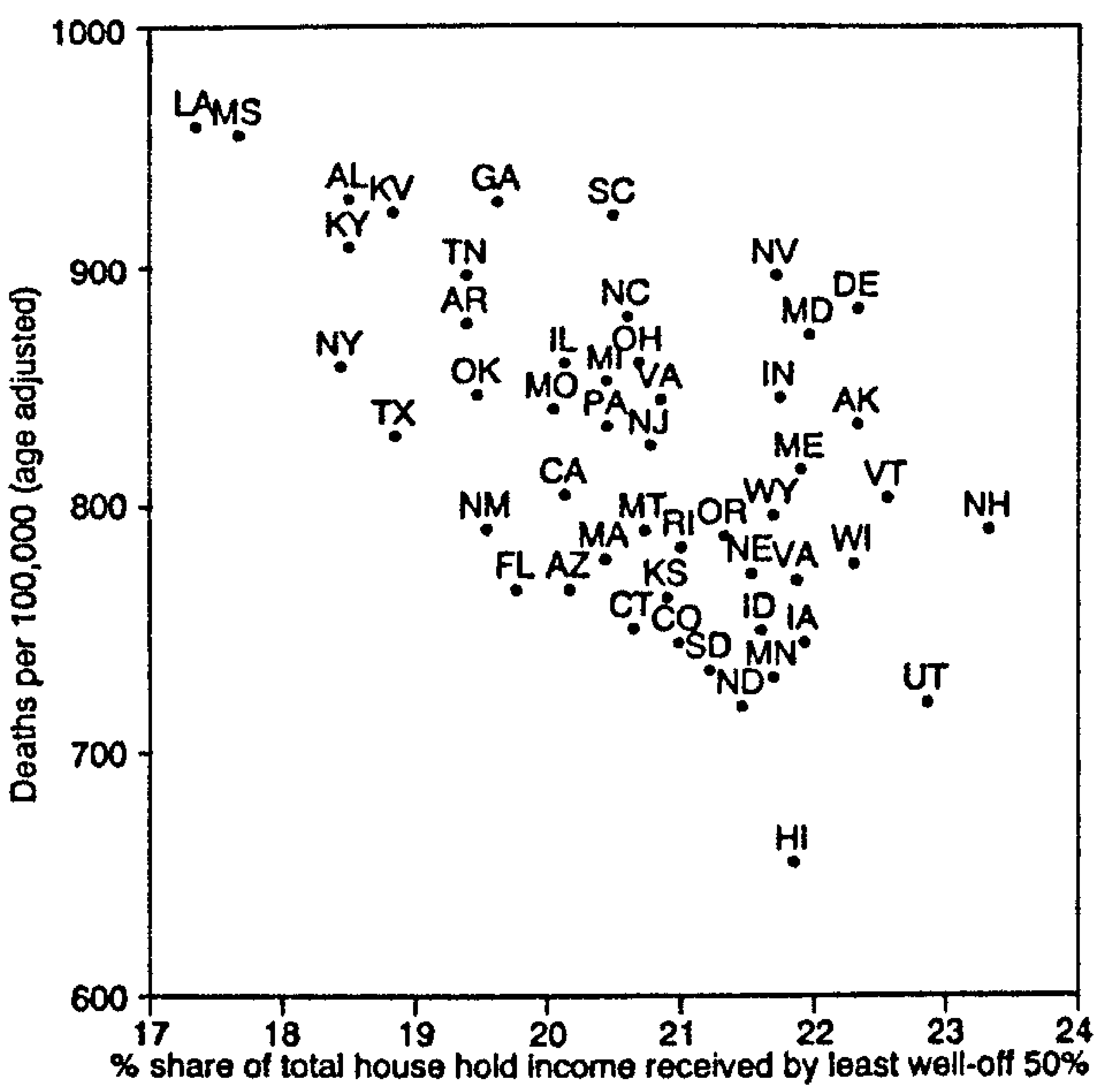
For Wilkinson the only explanation that accords with this evidence (relationship with absolute income between individuals within developed countries, but only with relative income between societies) is that the effects of incomes are not related to material disadvantage (poor quality housing, inadequate diet, poor access to facilities for sport and recreation) but to psycho-social processes. Living in a place with an unequal income distribution is anticipated to lead to a worse health experience. This is the effect of place or group, which is beyond the effect of the individual.

Figure 2.7 The annual rate of change of life expectancy in the twelve European Community countries and the rate of change in the percentage of the population in relative poverty, 1975-1985



Source: Wilkinson, (US, 1996, p77)

Figure 2.8 The relationship between income distribution and mortality among the fifty states of the USA in 1990



Source: Wilkinson, (US, 1996, p79)

"[w]ithin societies it is not so much the direct health effects of absolute material standards so much as the effect of social relativities... the relationship is with relative rather than absolute income levels. In the developed world, it is not the richest countries which have the best health, but the most egalitarian" (US, 1996, p3).

2.2.1.5 Evidence 5: towards social capital and social cohesion

Chapter 6 of Wilkinson's book *Unhealthy Society* provides a number of case studies to show how wider social structures are associated with narrower income differences and better health. These case studies are briefly been summarised as follows:

Example 1: Rosetto

This is a small town in Pennsylvania, USA, settled by Italian migrants from Rosetto. People who are residence here are much healthier than in neighbouring towns; Wilkinson argues that this is not due to reduced behavioural risk but an unusually close-knit community relationship with an egalitarian structure and high social cohesion. It was predicted that as Rosetto

"[l]ost its sense of community, it would also lose it health advantage - a prediction that a more recent study has borne out" (MTG, 2000, p13).

Example 2: Eastern Europe

Life expectancy in Eastern Europe and the Soviet Union were comparable with health in Western Europe during the 1960s. However, their worsening health since the early 1970s until 1989 was accompanied with

"[i]ncreasing inequality, by a growing air of cynicism in public life" (MTG, 2000, p14).

Again, for Wilkinson, as social cohesion breaks down, health and life expectancy worsen.

Example 3: Japan

In the early 1960s, Japanese life expectancy was lower than for Britain; however, it is now the highest in the World. To achieve parity with Japan, Britain would now have to have no deaths from all heart diseases and most cancers. The reason for this improvement in Japan, according to Wilkinson, is that *"[a] highly cohesive society - the culmination of a long period of narrowing income differentials and ... falling crimes"* (MTG, 2000, p14).

According to Wilkinson, the case studies above are by one major feature:

"[some] societies were unusually egalitarian and unusually healthy: circumstantial evidence suggests that they were also unusually cohesive" (US, 1996, p7).

2.2.1.6 Evidence 6: Income inequality and social disintegration

Wilkinson has found that

"[s]ocial disintegration is consequent on wider income differences" (US, 1996, p8).

He points to the study of Kaplan *et al.*, in 1990. They found a correlation of 0.72 ($P < 0.001$) between the homicide rate for forty-six states of the USA and the percentage of total household income received by the least well-off 50% of the population (Figure 2.9). This shows that

"[h]igher crime rates, homicide and violence are associated with wider income differences" (US, 1996, p8).

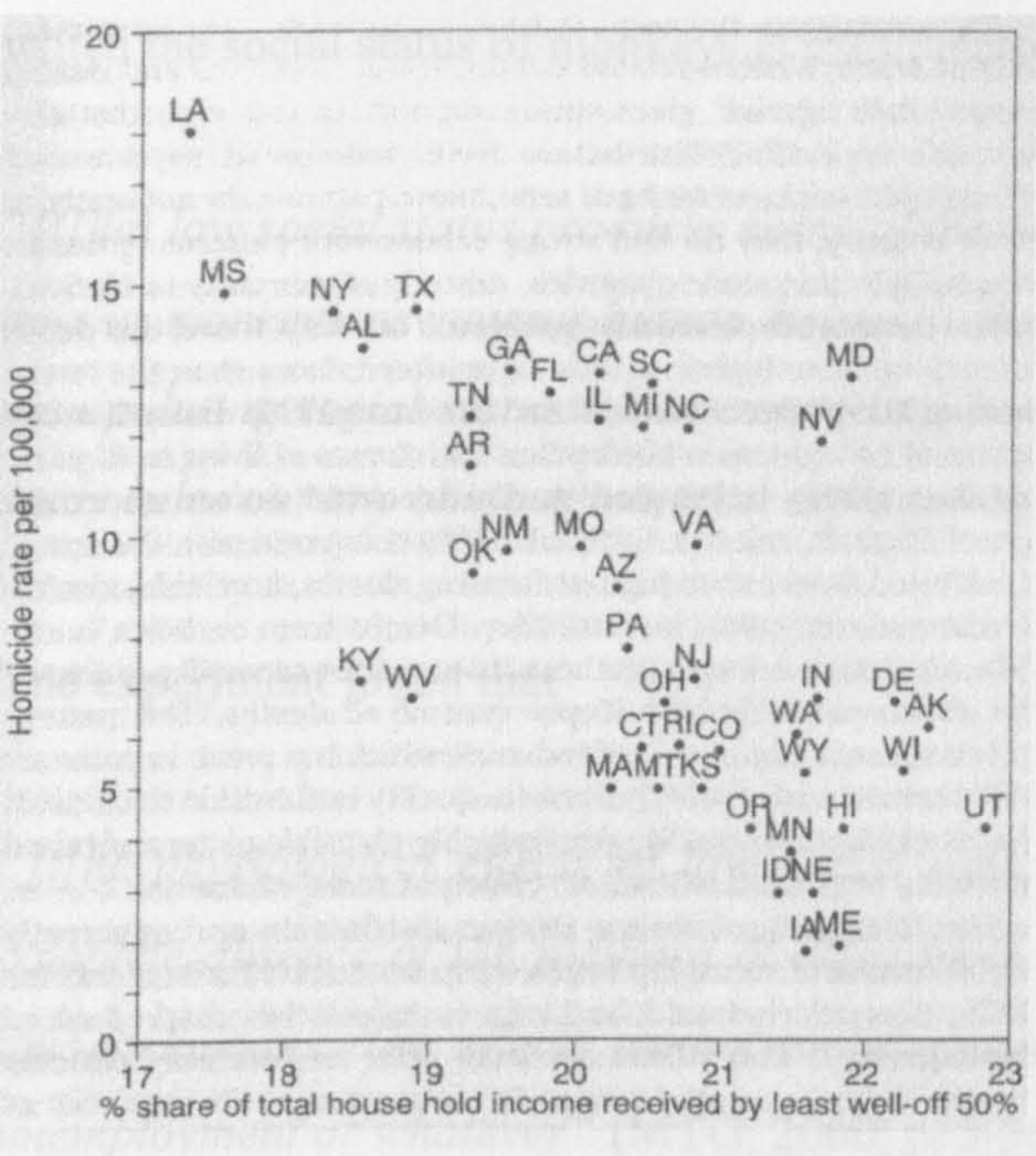
Wilkinson also puts emphasis on 'trust of others' and how this is related to income and health inequalities:

"[a]mong 50 States of the USA, the proportion of people who feel that they can trust others tends to be much higher in states where income differences are smaller. People in more egalitarian states find each other more helpful and are more likely to belong to voluntary clubs and associations. The causal relationship appears to run from income distribution through the quality of the social environment to health. There is also some international evidence that people trust each other more where income differences are smaller" (MTG, 2000, p14).

Also,

"[i]n terms of causal pathways affecting health, social cohesion may be regarded almost as an epi-phenomenon..... What affects health is the individual experience of chronic stress arising from social anxiety. Social cohesion is associated with health because it is a reflection of how tense or relaxed social contacts are and of the underlying social anxiety" (MTG, 2000, p59).

Figure 2.9 The relationship between income distribution and homicide among the states of the USA in 1990



Source: Wilkinson, (US, 1996, p157) after Kaplan *et al.*, (MBJ,1996, p1000)

2.2.1.7 Evidence 7: Psychosocial pathways

Wilkinson argues that the link between income inequality and health inequality is a result of psychosocial mechanisms, with anxiety as the key biological pathway:

"[i]ncreasingly it looks as if some of the most important parts of this relationship involve psychosocial pathways; subjective psychological and emotional effects of objective features of the social structure" (US, 1996, p54).

"[s]ocial structure and our position within it can exacerbate our anxieties about how we are seen by others- anxieties that go to the foundations of social life to our reflexive

nature as social beings" (MTG, 2000, p4).

"[i]n rich countries, anxiety seems to be one of the most important pathways linking health to social and economic circumstances" (MTG, 2000, p3).

Wilkinson bases his argument in part on animal studies. His book *Mind the Gap* was in a series about applying Darwin to modern society and evolution. For him there is increasing evidence showing that stressful social hierarchies predispose primates, such as troops of baboons to poorer health. Wilkinson reports the example of an experiment in which the social status of monkeys is manipulated, finding

"[t]hat low social status produces many of the same physiological risk factors for disease among monkeys as among human beings... the key processes ...are psycho-social; by manipulating social status while controlling both diet and the physical environment, the study rules out any obvious material explanation" (MTG, 2000, p12).

The experiment found that

"[t]he physiological effects of low social rank were observed in monkeys even in the complete absence of the plethora of differences in socio-economic circumstances found among humans. The effects cannot be attributed to jobs, housing, smoking, diet, debt, unemployment or whatever" (MTG, 2000, p35-36).

Moreover, it also has been found that

"[s]ubordinate monkeys and low-status humans share a much faster build up of atherosclerotic plaque in their coronary arteries, are more likely to suffer from central obesity, have potentially damaging levels of high-density blood fats..." (MTG, 2000, p36).

Turning now to human beings and, in particular, studies of workers, it has been found that

"[t]he amount of control people have over their work was found to be strongly predictive of health (even after controlling for occupational position and other positions) [...] a

series of studies showed that health worsened not only when people actually become unemployed but also before that when redundancies first seemed likely and people started worrying about their jobs" (MTG, 2000, p9).

Indeed, Wilkinson stresses the similarities of findings between primates and civil servants (the subjects of the Whitehall studies):

"[i]n monkeys the most common attacks are by dominant animals on subordinates. But higher fibrinogen levels are also much more common in junior staff, as if their subordinate positions put them at risk of psycho-social attack from their superiors" (MTG, 2000, p48).

Wilkinson puts the greatest explanatory power on the relative status of people in relation to people, rather than people in relation to things in developing his anti-materialist position

"[w]hat matters most about relative deprivation is not so much the lower living standards in themselves but the affront to dignity and respect, and the imputation of inferiority that accompanies relative poverty" (MTG, 2001, p27).

"[i]nstead of exposures to toxic materials and mechanical dangers we are discovering the toxicity of social circumstances and patterns of social disorganization" (US, 1996, p23).

2.2.1.8 Underlying causal model

It is possible to construct an underlying psycho-social model from a reading of Wilkinson's *Mind the Gap* and *Unhealthy Societies* as shown in Figure 2.10. He contends that the key mechanisms between income distribution and health operate through psycho-social processes. Countries where there is greater inequality, experience higher levels of social mistrust, higher social stress, greater unhappiness, poorer life satisfaction and consequently poorer health.

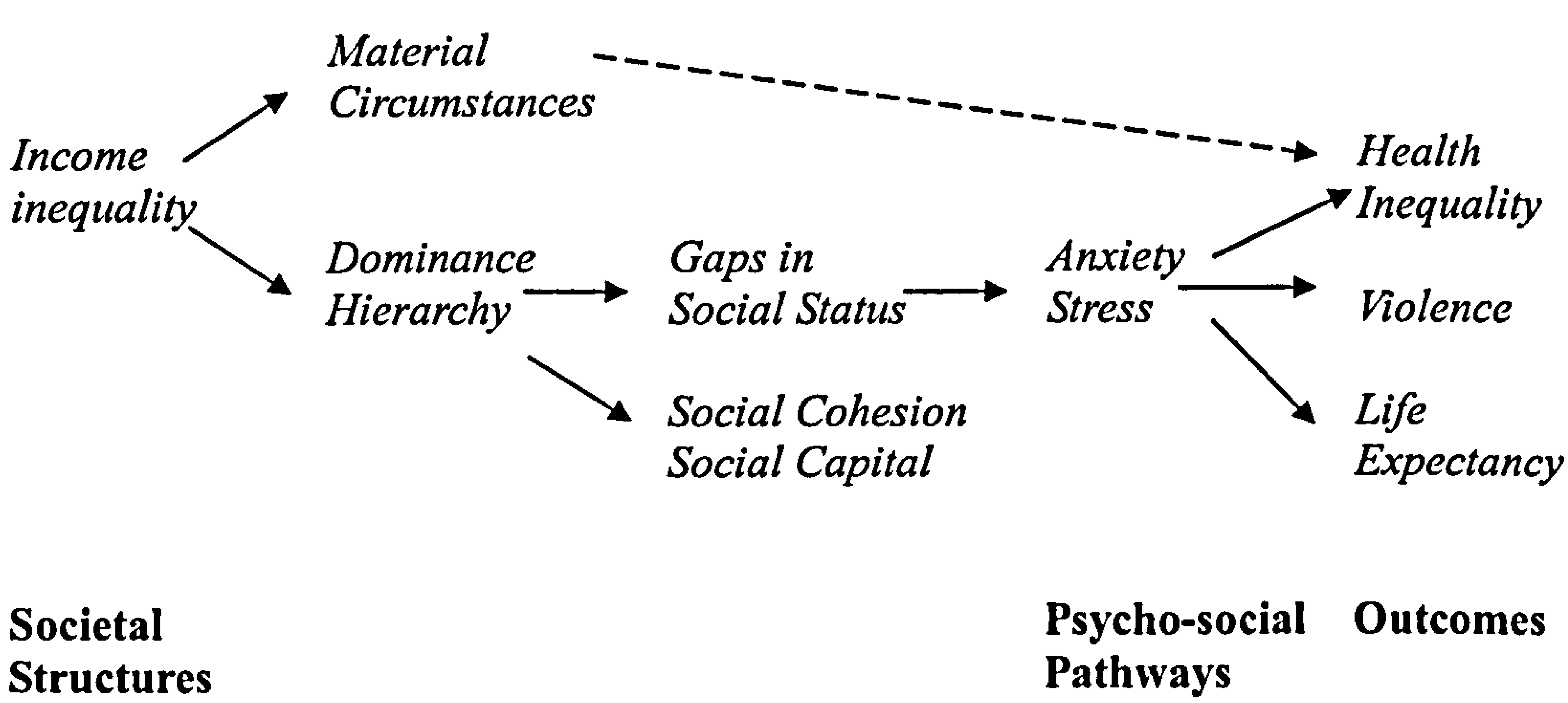
He argues that in developed economies, income inequality operates on health in only a minor way through poor material circumstances (such as inadequate diet and damp housing). For him, the key link is that unequal societies are socially divided societies with low social capital, poor civic engagement, and low levels of trust. This places individuals within such societies in a stressful situation, which then manifests itself as a range of 'social' problems, including violence, health inequalities and lower life expectancy:

"[m]ore equal societies are less stressful; people are more likely to trust each other and are less hostile and violent towards each other" (MTG, 2000, p3);

"[h]ealth and quality of life in modern societies as being primarily dependent on distributed justice and levels of what might be called social capital" (US, 1996, p9).

In summary, the Wilkinson hypothesis can be summarised in two statements. Firstly, in the developed world, the most egalitarian rather than the richest countries have the best health. Secondly, the most important links between disease and income inequality are psycho-social operating through the pathway of social cohesion

Figure 2.10 Underlying psycho-social model derived from a reading of Wilkinson's *Mind the Gap* and *Unhealthy Societies*



2.2.2 Critique

As befits an important and wider-ranging hypothesis, Wilkinson's arguments have been subjected to considerable criticism. Catalan (1998, 166) in a book review of *Unhealthy Societies* is sceptical of Wilkinson's very form of argument:

"[h]e marshals the work to support a conclusion he has already reached. He believes that economic disparity causes illness and implies there is no longer any empirical reason to think otherwise."

Catalan continues:

"[W]ilkinson's position that income disparities account for differences in life expectancy is actually based on something of a ***rhetorical device.*** *He describes two strong lines of research that are intuitively related to his argument and implies that the strength of this work compensates for the sparse findings of his own theory... I think it is fair to say that the empirical evidence for an association between income inequality and health is, as yet, suggestive. To imply, intentionally or otherwise, that the work is as compelling as that supporting the association between socio-economic status or social support and health is misleading."*

The Wilkinson argument has generated considerable debate and there are now literally hundreds of articles written in the past ten years supporting, developing and criticising his work (major reviews are provided by Gravelle, 1998; Kawachi and Kennedy, 1999; Wagstaff and Doorslaer, 2000; Macinko *et al.*, 2003 and Lynch *et al.*, 2004). Much of this interest derives from there being a great deal at stake in policy terms, between the competing demands for a policy aimed at material disadvantage and that aimed at social cohesion.

I will attempt to summarise this large body of criticism in terms of empirical critique, qualitative challenge, theoretical challenge and methodological challenge.

2.2.2.1 The empirical critique

A number of authors contend that Wilkinson's aggregated results are very sensitive to how inequality is measured, what countries are included, and the time period used. For example, Judge *et al.*, (1998, p567) found

"[v]ery little support for the view that income inequality is associated with variations in average levels of national health in rich industrial countries"

when using data from Australia, Belgium, Canada, Finland, France, Germany, Ireland, Italy, Luxembourg, the Netherlands, Norway, Sweden, Switzerland, the U.K., and the U.S.A, that is 15 countries with better data than that used by Wilkinson.

Lynch *et al.,* (2000), using aggregated data, plotted life expectancy published by the WHO against GDP (purchasing power parity, per capita) measured by the World Bank for 155 countries, as showed in Figure 2.11. They found that the relationship between absolute income and life expectancy does not plateau as Wilkinson claimed in Figure 2.2. Indeed, they found the correlation between life expectancy and GDP/head in this sample for the 33 wealthier countries to be r=0.51 (p=0.003). While this is shallower than the overall relationship between GDP/capita and life expectancy for all 155 countries, a correlation of this magnitude among wealthier countries would seem to qualify as evidence of a direct association. They point out that

"[i]n comparison with Wilkinson's selection of 23 countries, the addition of 10 equally wealthy nations that constitute the full sample significantly changes the results" (p405).

Consequently,

"[i]t is premature to dismiss the existence and importance of a direct association between absolute income and health status among developed nations" (p405).

In 2001, Lynch *et al.,* wrote a critique that revisited the question of whether income inequality is linked with health differences among wealthy economics. They pointed out that the evaluation of up to date wider-ranging international comparisons (of rich countries) on a variety of health outcomes has failed to find any significant correlation with income inequality. They used data provided by the Luxembourg Income Study (http://www.lisproject.org/techdoc.htm) for the period 1989-1991, and health outcomes such as low birth weight, life expectancy, self-rated health, and age- and cause-specific mortality in 16 countries.

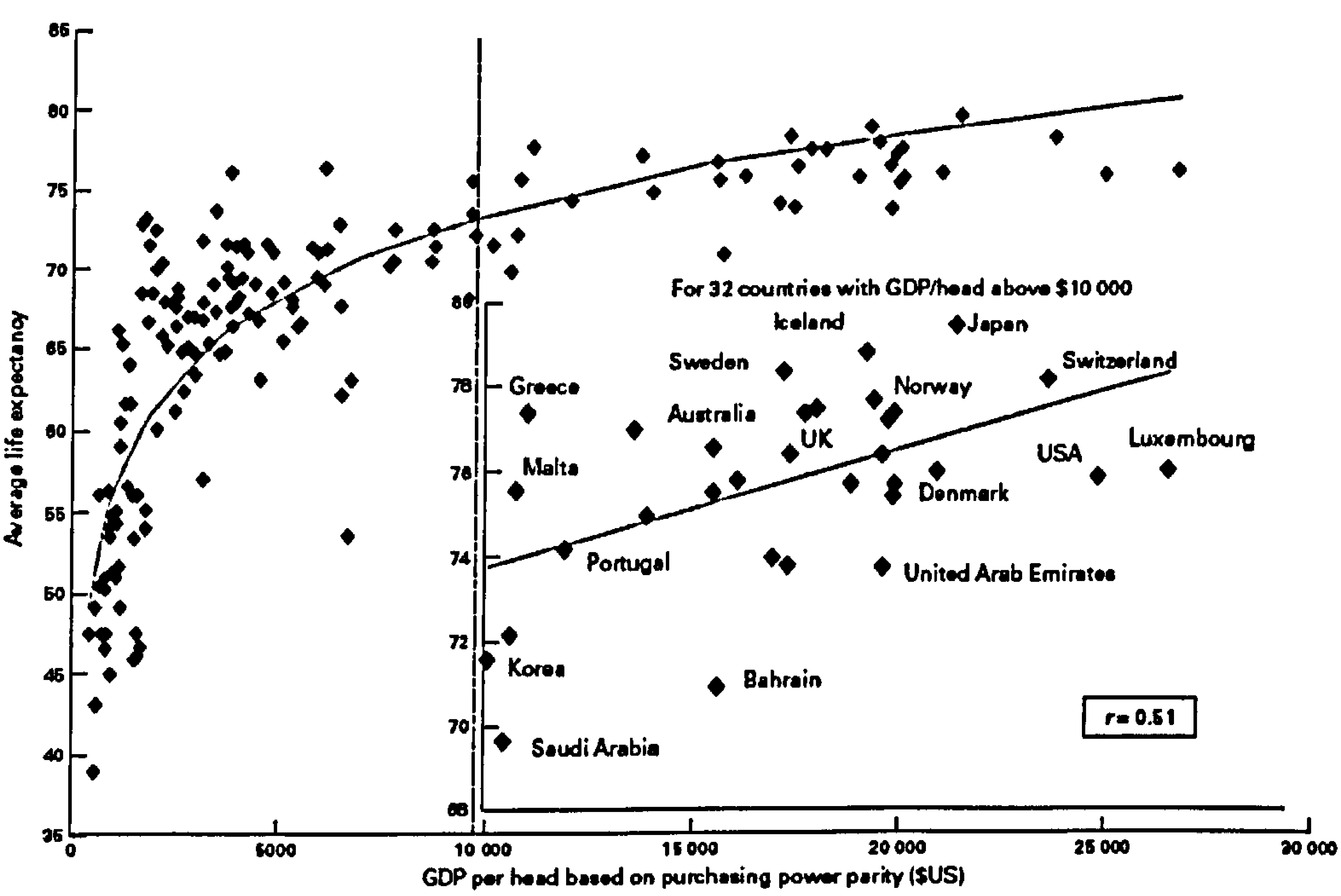
Figure 2.12a shows the results when Lynch and his colleagues use the same nine OECD countries as Wilkinson (Figure 2.5) but analyse the data for 1989-92. As Wilkinson found higher egalitarian countries (with lower income inequality) are associated with higher life expectancy (r = -0.45), although more weakly than in Wilkinson's original analysis (r = -0.81). After adding data for seven more developed countries (Italy, Spain, France, Belgium, Finland, Luxembourg and Denmark), Figure 2.12b shows that there is no longer a significant relationship between income inequality and life expectancy (r = -0.009, P = 0.75). It seems likely that if data from more countries had been available, Wilkinson's original study in 1992 would also have found little association between income inequality and life expectancy among the rich countries.

2.2.2.2 The qualitative / interpretivist challenge

Others have contended that Wilkinson work is too abstract and too generalizing in using aggregate data and correlation analysis. For example, Popay (2000) is doubtful about the debates between Wilkinson, Lynch and Judge, arguing that

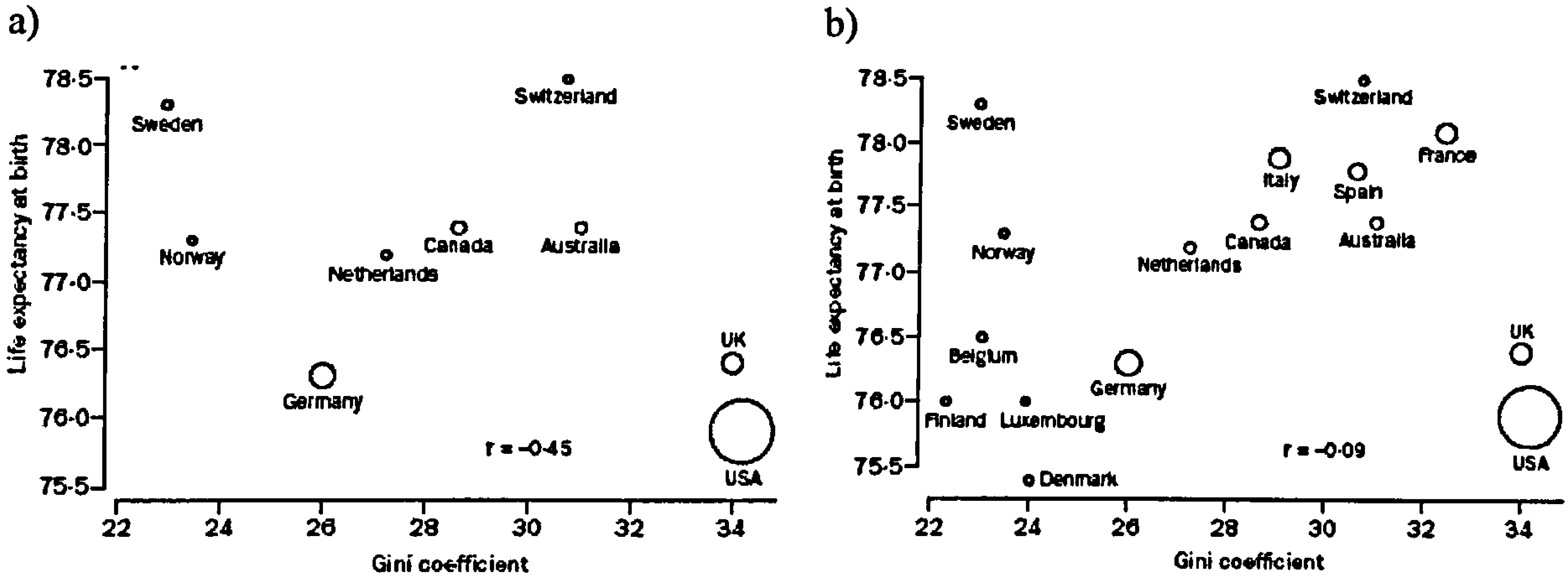
"[s]ocial capital must be conceptualized as a dynamic process involving people living in places... existing research fails to consider subjective, experiential dimension of social capital; it needs to move beyond the seemingly endless debates about macro- statistical relations...".

Figure 2.11 Association between GDP/head adjusted for $US Purchasing Power Parity and life expectancy for 155 countries circa 1993



Source: Lynch *et al.,* (JECH, 2000, p405)

Figure 2.12 Income inequality (1989-1992) and life expectancy at birth (1991-1993) (a) in the same nine countries used in Wilkinson (1992) in Figure 2.3 and (b) in the full sample of 16 countries



NB Circles represent country population size

Source: Lynch *et al.*, (Lancet, 2001, p198)

Working in the same qualitative tradition as Popay, Cattell (2001) undertook an in-depth study of two places in London and concluded that

"[i]n Wilkinson's, use of social cohesion and social capital ... mechanisms such as shame, disrespect social anxiety and perceptions of inferiority are induced by interacting with people of higher social status. Such mechanisms might well occur in certain situations – such as the workplace- but this research suggest that these processes may be less useful an explanation for what happens in poor areas... Conversely strong perceptions of inequality residents may be motivated to take cooperative action which carries the potential for benefiting health" (p1513).

She continues

"[a]s a concept which bridges structural and cultural approaches to poverty, social capital... is not wholly adequate for explaining the deleterious effects of poverty on health" (p1514).

2.2.2.3 The political/theoretical challenge

Another set of challenges has come from the ideological left. In 2000, Coburn thought

Wilkinson ignores or downplays the causes of inequality and argued that it must be seen as an outcome of neo-liberal market ideology, changing class structure and globalisation. That is he argues that income inequality is not the root cause of health inequalities, but is an intervening variable with the root explanation being the working of the global capitalist system. Muntaner (1999) and Muntaner *et al.,* (1999) develop a neo-materialist response arguing that Wilkinson psychopathologizes the relatively deprived, omits social determinants of disease related to production and indulges in 'community blaming' whereby an erroneous characterization of working-class communities as non-cohesive could be used to justify paternalistic or punitive social policies. Lynch and his colleagues (2000) use the air-travel metaphor to make the neo-materialist argument in a direct way. They ask why do second- class travellers get backache on long-haul travels. They dismiss the psychosocial interpretation that its is due to second-class travellers being able to see the better condition found in first class, contending that the real cause is the materialist cramped conditions and inadequate seating.

2.2.2.4 The methodological challenge

From reading this literature (e.g. Gravelle, 1998; Wagstaff and van Doorslaer, 2000) there are four important methodological points to be made. Firstly, much of the literature involves relatively simple analysis in which a single measure of income is plotted against a measure of health outcome (as in Figures 2.1 to 2.9): this is problematic as it would be extremely useful to know if the relationship between mortality and inequality holds when account is taken of absolute income and *vice versa*. Secondly, much of the literature uses correlation and not regression. Correlation measures the scatter around the line and not the size of the effect as would a regression coefficient.

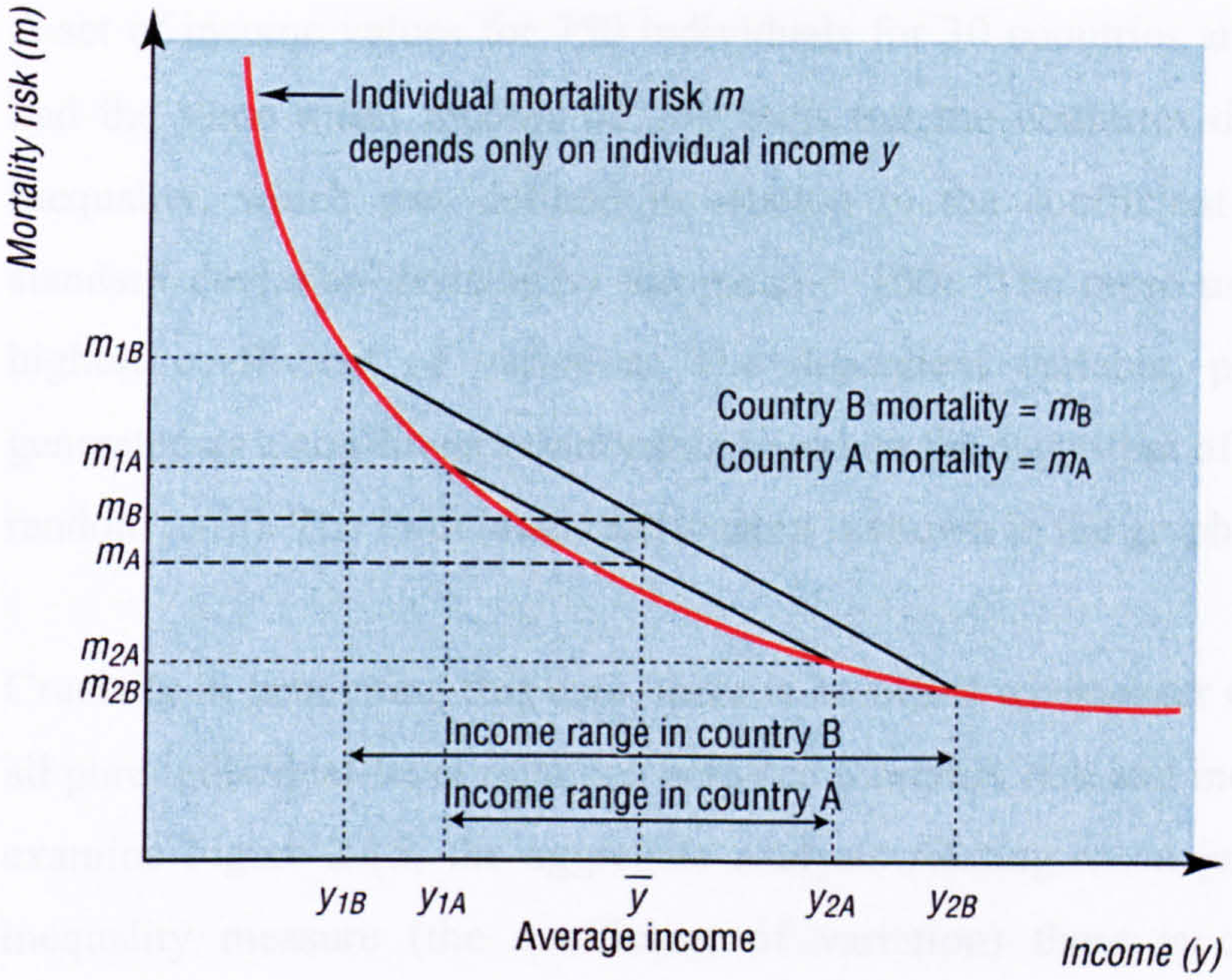
Thirdly, much of this literature is cross-sectional and does not consider change over time. This is problematic as we should be concerned with transitions: the case for a causal relationship is made most powerfully when a change in the cause, be it between income or inequality, precedes a change in the outcome of life expectancy and transition. Finally, much of this literature is based on aggregate analysis where mean income or inequality of a country is related to an overall mean death rate. It is this problem of aggregate analysis which is now considered in more depth, preceding the distinctive focus of this thesis.

Wilkinson's empirical evidence for the income inequality hypothesis is largely based on the result of aggregated data analysis on developed countries. Such an approach can only infer relations at the individual level from associations between variables at the population level, and might be subject to an ecology fallacy (Rodgers, 1979). That is that at least part, and possibly all of the population mortality associated with inequality of population level measures, are biased estimates even after conditioning for other potentially important determinants of mortality. In the classical analysis of the ecological fallacy, Robinson (1950) found that in a "means-on-means" analysis that aggregate results for USA States did not match those found for individuals.

Gravelle (1998) gives the argument a new twist by showing that when there is a non-linear relationship at the individual level, an aggregate analysis will shows an artefactual relation between the mean of the response and the variance (or inequality) of the predictor. In the present case, if there is a non-linear relationship between the individual health outcome and individual income, the importance of any relative income effect will be a "statistical artefact".

This is clearly explained by Gravelle's example in Figure 2.13. This picture is under the assumption that individual mortality risk (m) is only determined by personal income (y), and that there are no other effects, such as income distribution and average income. The underlying true relation is a strong non-linear relation between income and health. It is a simple example where there are only two countries, A and B, having the same average income ($\bar{y}$) but different degrees of income distribution such that country B is more unequal than A (the range y_{2B}-y_{1B} is greater than y_{2A}-y_{1A}). Country B is a higher mortality risk (mB = m1B-m2B: the differences of the function between lowest and highest individual income). Under such a non-linear relationship between individual income and health outcomes, transferring income from the rich to the poor will help to decrease the mortality risk of the poor more than it increases the mortality of the rich. In general, as income inequality between countries reduces, population mortality declines even though every individual's risk of mortality only depends on their own income level and not on the income level of anyone else.

Figure 2.13 Effect of increased inequality of income on population mortality



Source: Gravelle, (BMJ, 1998, p383)

2.3 THE NEED FOR A MULTILEVEL METHODOLOGY

To re-iterate, the key crux of Wilkinson argument is that

"[w]here income is related to social status, as it is within countries, it is also related to health. Where income differences mean little or nothing for people's position in the social hierarchy (such as those between countries), income makes little difference to health. This strongly implies that psychosocial pathways are important" (MTG, 2000, p10-11).

But as we have just considered, for Gravelle (1998) there is nothing surprising in that there is a relationship between income inequality and mean mortality; it is purely an artefact of the aggregation process. However, Gravelle does not provide a way of transcending the ecological fallacy as for him the analysis is either individual or it is aggregate. Here I want to build on his insight by showing that individual data are necessary to distinguish between the absolute and relative income hypothesis and that a methodology is required that analyses regression-like relationships simultaneously at the individual and place level: the so-called multilevel approach. In this section I simulate data with known functional relations and then analyse these data firstly as an aggregate (and incorrect) analysis in the manner of Wilkinson and then as a multilevel model.

2.3.1 Simulation study and aggregate analysis

A set of income values for 750 individuals for 30 countries are simulated. All countries had the same mean income of 750 units but the countries differed in their degree of inequality, which was defined in relation to the coefficient of variation (that is the standard deviation divided by the mean * 100). The more unequal countries have the highest coefficient of variation. The dependent variable, probability of death, was generated as a non-linear relationship based on the logarithm of individual income (plus a random term). The individual relationship is shown in the graph of Figure 2.14.

Crucially in generating this data, there is no effect whatsoever of country inequality; it is all pure individual-level relations between mortality risk and income. However, when we examine Figure 2.15, the aggregate analysis relating mean probability of death to the inequality measure (the coefficient of variation) there is clearly a positive strong relationship, which is a complete artefact of the aggregation process.

Figure 2.14 Simulated non-linear relation between income and mortality

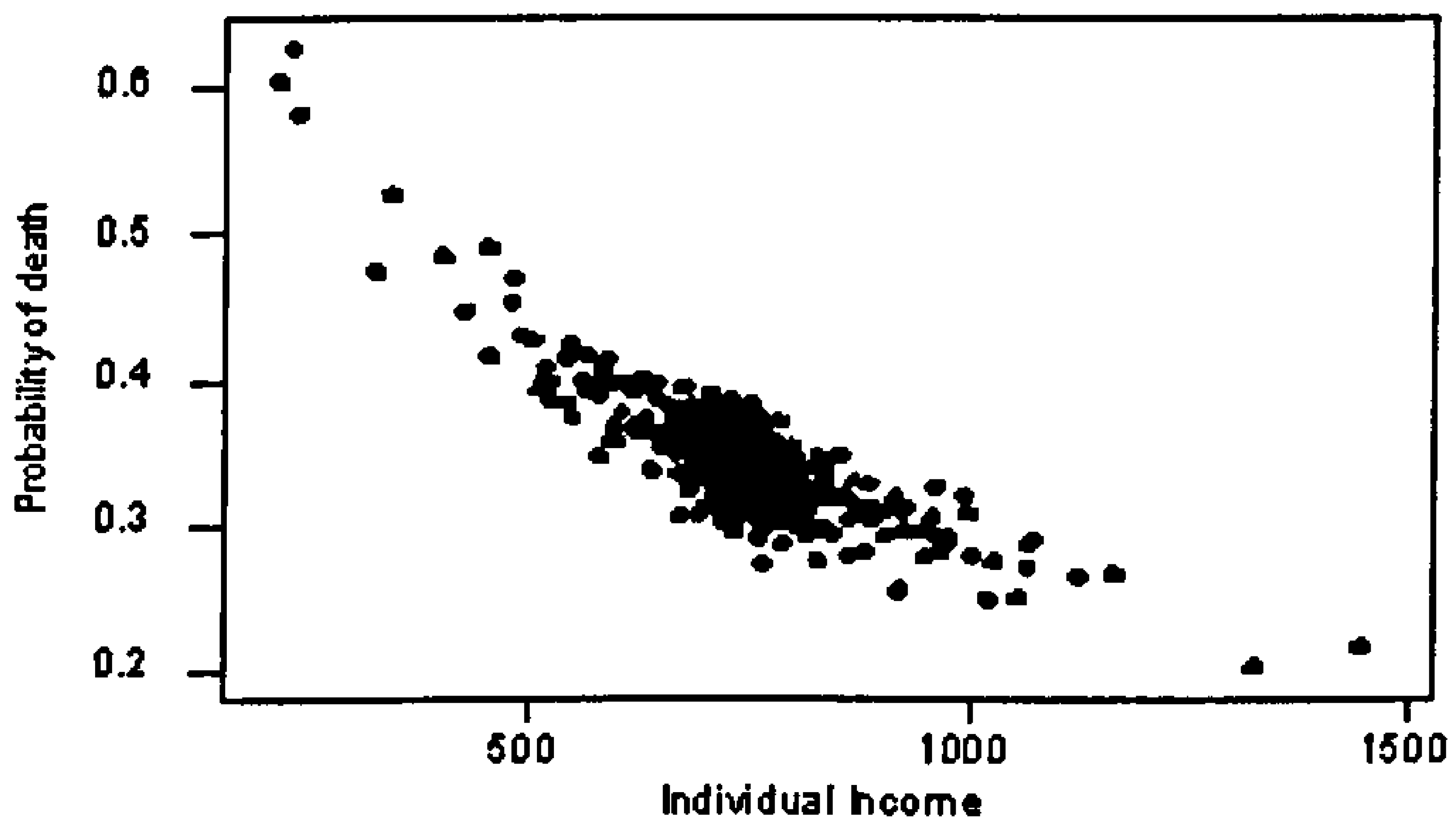
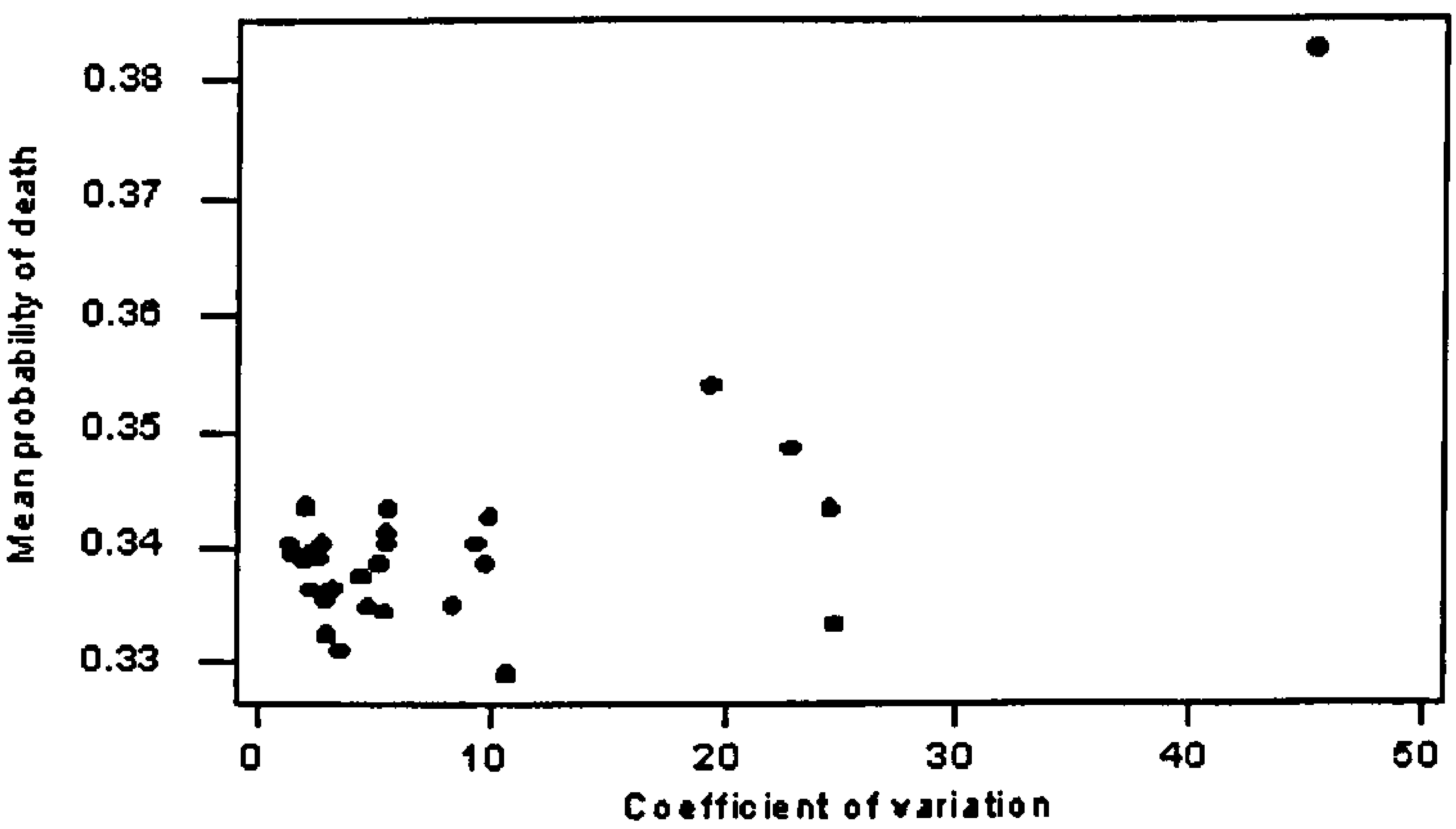


Figure 2.15 Artificial relationship between mortality and inequality



Unequal countries look to be unhealthy places but this is an outcome of an aggregate analysis of an underling non-linear individual relationship. How this comes about is made clear by the next graph (Figure 2.16), where we look at the extremes of countries 30 and 29, which are the most and least equal countries, with blue and red symbols respectively. The 'blue' country has a greater number of people on low income, who because of the non-linearity have a particularly high probability of death, but this is not 'counterbalanced' by a particularly low mortality for those who individually have high income. The 'red' country has its people located in the centre of the distribution, and there are no particularly poor people with their associated high death rates.

2.3.2 Multilevel modelling

The above description of the simulated data makes it clear that we are dealing with data with a complex structure of people (individuals) and places. Individual income is needed and we need to distinguish between this variable and ecological (place-level) variables of mean income and income inequality. Conceptually we can think of a range of results so that only individual income affects health; or ecological characteristics affect health in addition to individual effects; as that individual and ecological effects work in interaction with each other, so that for example poor people in poor places have particularly poor health experience. We need to test empirically these rival hypotheses, and the appropriate quantitative methods are multilevel models, which provide estimates for data with a complex hierarchical structure (Goldstein, 2002). Moreover, only individual-level data has the potential to discriminate between such hypotheses.

To illustrate the structure of a multilevel model, I turn now to a multilevel analysis of the simulation analysed incorrectly above. I will begin by fitting a two level null model in which probability of death (multiplied by 100) is estimated as an overall average and the remaining variation is appear decomposed to between-country and within-country variance. It is called a null model because there are no predictor variables except the constant and the associated intercept. The simplest possible two-level null model is specified as follows:

$$Y_{ij} = \beta_0 X_{0ij} + (\mu_{0j} X_0 + \varepsilon_{0ij} X_0) \qquad (2.1)$$

where

Y_{ij} is the probability of death (*100) for person i in country j;

X_{0ij} is a constant, just a set of 1s for each and every individual in every country;

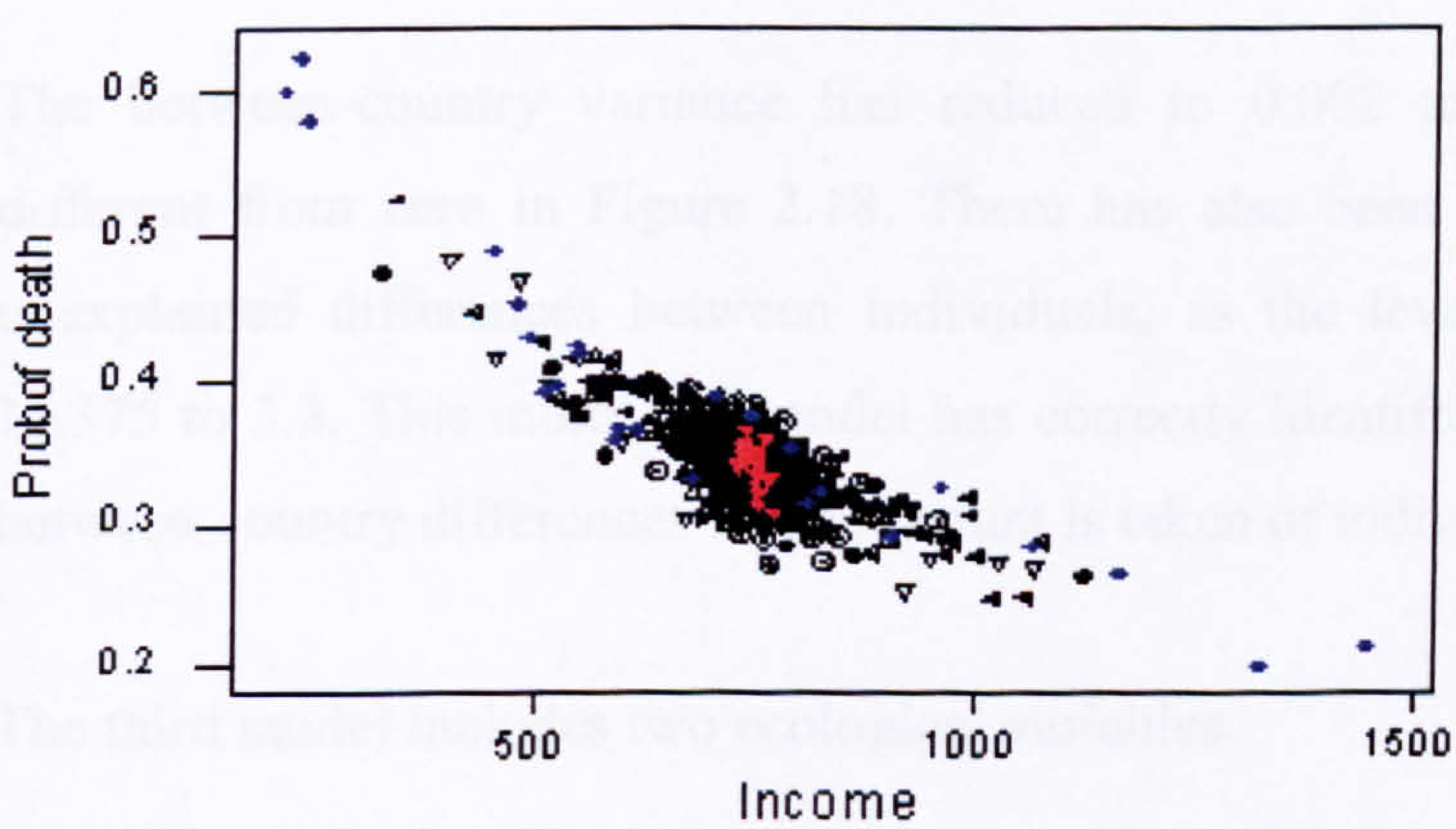
β_0 is the overall mean probability of death across all countries;

μ_{0j} is the differential in mortality risk for being in country j, from the global mean-these are known as the random differential intercepts; and

ε_{0ij} is the differential in mortality risk for being person i in country j.

The μ_{0j} are assumed to come from a Normal distribution and to be summarised by the variance σ^2_{u0}, while the ε_{0ij} are also assumed to come from a Normal distribution with a variance of σ^2_{e0}. Figure 2.17 shows the results of fitting this model in MLwiN (Rasbash *et al.*, 2004) to the simulated data. The mean probability of death×100 is 34, and around this global average there are variances of 0.407 for between country variation (σ^2_{u0}) and 11.375 for within country, between-person variability (σ^2_{e0}). The figures in brackets are the estimated standard errors which can be used to place approximate confidence intervals on the parameters.

Figure 2.16 Identifying least and most unequal countries indicators for countries 29 and 30



Using a Wald test (Goldstein, 2003) the between-country variance of 0.407 is significantly different from zero ($p < 0.10$). The main question is whether this is a true country effect or is it due to the nature of the people within a country, which is a compositional effect.

Figure 2.17 Results for a two-level null random intercepts model for simulated data

$$\text{Probdeath}_{ij} \sim \text{N}(XB, \Omega)$$
$$\text{Probdeath}_{ij} = \beta_{0ij}\text{cons}$$
$$\beta_{0ij} = 34.016(0.170) + u_{0j} + e_{0ij}$$

$$\left[u_{0j}\right] \sim \text{N}(0,\ \Omega_u) \ : \ \Omega_u = \left[0.407(0.224)\right]$$

$$\left[e_{0ij}\right] \sim \text{N}(0,\ \Omega_e) \ : \ \Omega_e = \left[11.375(0.600)\right]$$

*-2*loglikelihood(IGLS Deviance)* = 3971.129(750 of 750 cases in use)

The model is now developed to include log of income for individuals as a predictor, seeing what happens to the fit of the model and the size of the between country effect:

$$Y_{ij} = \beta_0 X_{0ij} + \beta_1 X_{1ij} + (\mu_{0j} X_0 + \varepsilon_{0ij} X_0) \qquad (2.2)$$

where X_{1ij} is the log of the income of person i in country j centred around the global mean across all countries and all individuals;

and β_1 is the overall global relation between mortality and individual income.

The between-country variance has reduced to 0.002 and is no longer significantly different from zero in Figure 2.18. There has also been a substantial reduction in the unexplained differences between individuals, as the level 1 variance has fallen from 11.375 to 3.3. This multilevel model has correctly identified that there are no significant between country differences when account is taken of individual income.

The third model includes two ecological variables

$$Y_{ij} = \beta_0 X_{0ij} + \beta_1 X_{1ij} + \alpha_1(W_{1j}) + \alpha_2(W_{2j}) + (\mu_{0j} X_0 + \varepsilon_{0ij} X_0) \qquad (2.3)$$

where W_{1j} is the mean income in country j centred around the global mean across all countries having taken account of the individual income-health relation;

W_{2j} is the level of income inequality in country j centred around the global mean across all countries having taken account of the individual income-health relation;

and α_1 estimates the effect of mean country income on individual health
α_2 estimates the effect of country's income inequality on individual health

The estimated parameters for mean country income and country inequality are correctly found not to be significant (Figure 2.19). There are between-place effects in the null model, but it goes away when individual income is included. While the aggregate analysis cannot distinguish the effect of individual income and income inequality, the multilevel model can.

Figure 2.18 Results for a two-level random intercepts model for simulated data, after inclusion of log of individual income

$$\text{Probdeath}_{ij} \sim \text{N}(XB, \Omega)$$
$$\text{Probdeath}_{ij} = \beta_{0ij}\text{cons} + \text{-}20.752(0.473)\text{Loginc-6}_{ij}$$
$$\beta_{0ij} = 46.583(0.294) + u_{0j} + e_{0ij}$$
$$\left[u_{0j}\right] \sim \text{N}(0,\ \Omega_u) : \Omega_u = \left[0.002(0.035)\right]$$
$$\left[e_{0ij}\right] \sim \text{N}(0,\ \Omega_e) : \Omega_e = \left[3.300(0.174)\right]$$

*-2*loglikelihood(IGLS Deviance)* = 3024.323(750 of 750 cases in use)

Figure 2.19 Results for a two-level random intercepts model for simulated data, after inclusion of log of individual income, mean country income and country inequality

$$\text{Probdeath}_{ij} \sim \text{N}(XB, \Omega)$$
$$\text{Probdeath}_{ij} = \beta_{0ij}\text{cons} + \text{-}20.524(0.481)\text{Loginc-6}_{ij} + \text{-}0.006(0.005)\text{MeanInc}_j + 0.010(0.008)\text{CVAR}_j$$
$$\beta_{0ij} = 50.827(3.576) + u_{0j} + e_{0ij}$$
$$\left[u_{0j}\right] \sim \text{N}(0,\ \Omega_u) : \Omega_u = \left[0.000(0.000)\right]$$
$$\left[e_{0ij}\right] \sim \text{N}(0,\ \Omega_e) : \Omega_e = \left[3.276(0.169)\right]$$

*-2*loglikelihood(IGLS Deviance)* = 3018.481(750 of 750 cases in use)

2.3.3 Multilevel studies

It is clear from the above methodological discussion that a multilevel approach will allow a proper evaluation of the Wilkinson hypothesis. Table 2.1 presents the published studies that have used multilevel analyses to investigate the relationship between area-level income inequality and individual-level health outcomes, taking into account certain individual-level socio-economic predictors (e.g., individual income). This table has been updated to December 2005 from Subramanian's paper (2002), omitting studies where a full multilevel methodology has not been adopted[2]. The papers have been found by defining "multilevel studies" as those that utilize multilevel data in the form of an individual-level health outcome, a set of individual-level socioeconomic predictors (e.g., individual income), and an area-level income inequality measure (e.g., State income inequality).

The details of the setting, method and findings are given in summary form in the table, but I would like to emphasize three important issues. Firstly, except from Elgar's study in 2005, no other study uses multiple countries and individual-level health outcome data to evaluate Wilkinson's hypothesis at the scale at which he developed it; for him 'societies' in his relative income hypothesis is defined operationally as countries. Secondly, no study examines the hypothesis over an extended period of time. Thirdly, eight out of thirteen studies provide support for the income-inequality hypothesis but rest of the studies that suggest income inequality is associated with poor health have been carried out within the United States. Studies conducted outside the United States have generally failed to find an association between income inequality and health.

A more recent trend is the attempt to unpack the mechanisms involved between income inequality and health outcomes, particularly by the mediating process of social cohesion. For example, one of the earliest studies, Kawachi *et al.,* (1997) examined the links between social capital, defined as per capita density of membership in voluntary groups in each state, and levels of social trust as a pathway between income inequality and population health.

---

2 Some studies have used 'marginal' models in which the standard errors are corrected for within-place similarity, but a full multilevel approach is not taken that allows for the estimation of between-place differences.

Table 2.1 Empirical evaluation of the Wilkinson's income inequality hypothesis using individual data (under multilevel modelling analysis)

Author (Date)	Main outcome	Income inequality Measure	Average Income Measures	Study design, Data set, Population, sample size	Income Inequality Related to Health?	Study Conclusions
Elgar et al. (2005)	Self-rated data on alcohol consumption and episodes of drunkenness	Gini index	Family income; was measured using the HBSC Family Affluence Scale (FAS) Country wealth: GDP per capita	Cross-sectional study, The health behaviour in school-aged children study surveyed in 34 countries, 11, 13 and 15 adolescents. N= 162,305	Weak yes	Income inequality may have a contextual influence on the use of alcohol among younger adolescents. Findings suggest that economic policies that affect the distribution of wealth within societies may indirectly influence the use of alcohol during early and mid-adolescence.
Xi et al. (2005)	Self-rated health	Individual health status associated with income inequality (Gini coefficient) at the public health unit level	Median area income	Cross-sectional study, Ontario Health Survey in 1996-97. 25+ residences in Ontario, N=30,939	Yes	Income inequality was significantly associated with individual self-reported health status at public health unit level in Ontario, independent of individual income.
Wen et al. (2005)	Mortality	Individual income status (poverty), neighbourhood income inequality	The percentage of residents with household annual incomes $50,000 or over (concentrated affluence), The percentage of households in a neighborhood that were below the Federal poverty threshold in 1990 - $13,359 for a household of four (concentrated poverty)	Cross-sectional study, A comprehensive social survey of neighbourhood residents who are survivals after hospitalisation in 1994-95. N=10,557	No	Community-level social interventions based on social capital/social cohesion models are not likely to achieve fruitful results without concomitant effort in the economic and health care realm, at least in terms of influences on the health of older people once they are already ill.
Henderson et al. (2004)	Alcohol dependence, depression	Gini	Median income	Cross-sectional study, US national probability sample US residences,	No	No association between income inequality and alcohol dependence or depression after control for individual covariates.

				N=42,862		
Jones, Duncan, & Twigg (2004)	Mortality	Robin-Hood Index (HALS) and Gini (HALS) at parliamentary constituency level	Mean income, mean income (HALS)	Longitudinal study, British Household study UK residences, N=8720	No	When account is taken of two socio-structural variables, and four health-related behaviours, area income in the form of Parliamentary Constituency mean income, remains a significant contributor to the variations in individual mortality.
Blakely, Atkinson, & O'Dea (2003)	All-cause and cause-specific mortality Sex-specific mortality	Gini	Mean household income	Three years follow up mortality study, New Zealand residences aged 25-64, N=1,391,121	No	No association between income inequality and mortality after adjustment for individual covariates. Not associated with any specific cause of death.
Subramanian & Kawachi (2003)	Self-rated health	Gini	Median income	Cross-sectional study, Current Population Survey, 1995, 1997, US adults, N= 201,221	Yes	A 0.05 increase in income inequality after adjustment for individual-level factors, associated with approximately 30% higher odds of poorer self-rated health.
Subramanian, Blakely, & Kawachi (2003)	Self-rated health	Gini	Median income	Cross-sectional study, Current Population Survey, 1995, 1997, US adults 45+, N= 90,000	Yes	A 0.05 increase in income inequality, after adjustment for individual-level factors, associated with approximately 15 to 18% increased odds of poorer self-rated health
Subramanian, Blakely, & Kawachi (2003)	Self-rated health	Gini	Median income	Cross-sectional study, Chile adults aged 15+, N=98,344	Yes	Comparing four categories, higher income inequality associated with approximately 20% increased odds of reporting poor self-rated health
Blakely et al. (2002)	Self-rated health	Gini	Average household income	Cross-sectional study, Current Population Survey, 1995, 1997, US adults, N= 18,547	No	There is only a small association of MA-level income inequality with fair/poor health when controlling further for average MA household income. Regarding the association of state-level income inequality with fair/poor health, the association has been found to be considerably stronger among non-metropolitan (i.e. rural) compared to metropolitan residents.

Subramanian, Kawachi, & Kennedy (2001)	Self-rated health	Gini at state level	Median income at state-level	Cross sectional study, Behavioural Risk Factor Surveillance Study 1993-1994, US adults, N= 144,692	Yes	Strength of overall association not reported, but higher income inequality associated with poorer self-rated health, but only among higher-income groups. Not associated with self-rated health among middle- and low-income groups.
Blakely et al. (2000)	Self-rated health	Gini (state-level)	Median household income, state income	Cross-sectional study, US Current Population Survey, 1995 1997, US residents aged 15+, N=213,695	Yes	A 0.05 increase in the Gini associated with 30% increased odds of poorer self-rated health but effects observed only among those aged 45+, and this effect may be lagged by 15 years.
Diez-Roux et al. (2000)	Sex-specific prevalence of smoking, Body Mass Index, hypertension and sedentarism	Gini, Robin Hood Index, 50% share (by state)		Cross-sectional study, Behaviour Risk Survey Surveillance System, Us adults in 50 states, N=81,557	Yes	For men, higher income inequality associated with approximately two-fold increased odds of sedantarism, but not associated with smoking, hypertension, or higher body mass index. For women, higher income inequality associated with increased odds of sedantarism; 61% increased odds of hypertension, but only among poorer women, whereas income inequality associated with smoking but only among wealthier women.

Figure 2.20 shows rapid and continuing increases in the number of papers mentioning 'social capital' since 1998 and this has now reached 75 publications a year in 2005.

Figure 2.20 "Social capital" indexed papers in MEDLINE, 1992-2005

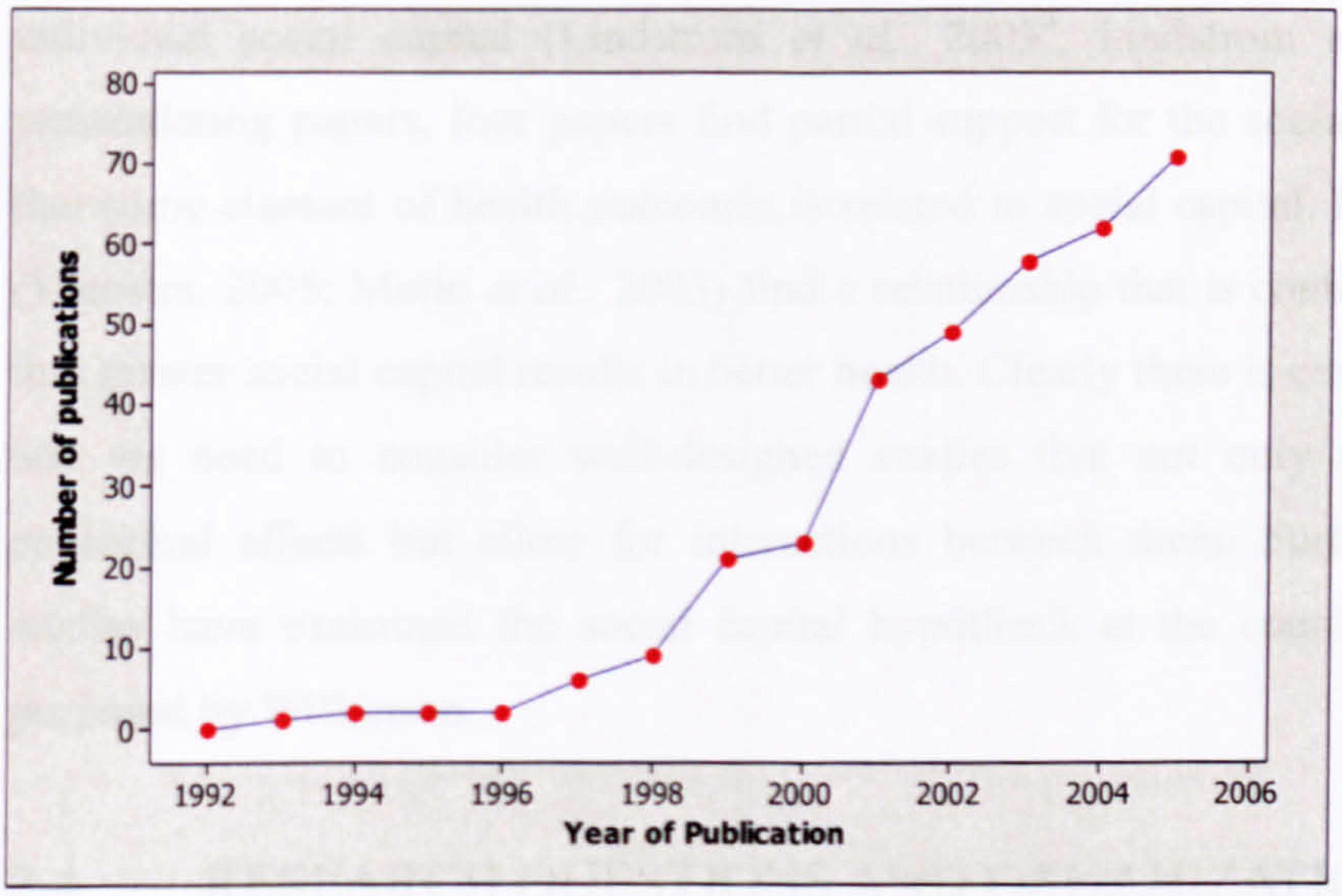


Table 2.2 summaries the existing 16 empirical studies that have applied full multilevel analysis to social capital influences on population health. The table has been produced by updating until the end of 2005 from Kawachi's review and commentary of 2004. The table illustrates the diversity in the choice of indicators used to measure social capital across published studies. Some combination of measures of trust, perceived reciprocity, and social participation are common in most studies and then these are aggregated to the community or State level. Others measures of social cohesion include volunteerism, community attachment and electoral participation. Given this diversity of measures of social capital and cohesion, it is perhaps unsurprising that there is no consistent relationship with health.

The table identifies two types of study

1. C only: where C refers to a contextual or ecological variable such as the percent of population within a place that are active volunteers; and
2. C and I: where I refers to social capital variables defined at the individual level, such as whether the respondent is an active volunteer or not.

The nine studies that use only ecological ('C') variables are methodologically problematic as they cannot separate out composition (what is in a place) from context (the difference a place makes). Seven out of the sixteen studies research contextual and compositional effects simultaneously. One study finds that both are important (Subramanian *et al.*, 2002). Two other studies only found significant relationship between individuals' health and individual social capital (Lindstrom *et al.*, 2003[a], Lindstrom *et al.*, 2003[b]). Of the remaindering papers, four papers find partial support for the social capital hypothesis in that some element of health outcomes is related to social capital. However, two of these (Veenstra, 2005; Merlo *et al.*, 2003) find a relationship that is contrary to the relationship that greater social capital results in better health. Clearly there is conflicting evidence here and we need to consider well-designed studies that not only model individual and contextual effects but allow for interactions between them. Furthermore, none of the studies have examined the social capital hypothesis at the country level as originally purposed by Wilkinson.

2.4 RESEARCH QUESTIONS AND ORGANIZATION OF THESIS

Having reviewed the Wilkinson hypothesis, made the argument for a multilevel approach and summarised previous multilevel studies this final section of the chapter consider the aims of the rest of this thesis. The primary aim of this thesis is to evaluate empirically the Wilkinson argument that the major determinant of health outcome in countries with advanced economies are relative inequalities in income, and to do so taking account of different sub-populations (based on gender, age and income and allowing for the possibility that different countries may experience the effects differently.

To that end the study is based on four distinct pieces of research, which are all international-based comparisons:

1 A global study of changing aggregate mortality patterns: including an examination of overall life expectancy for 169 countries. The aim of this analysis is to uncover the basic trends in mortality between countries using comparable data and to relate these to changing levels of income and income inequality. This is study is the subject of Chapter Three.

Table 2.2 Multi-level studies of social capital and health

Author (Date)	Main outcome	Measure of social capital[3]	Sample size, Data structure, Data set	Study Conclusions
Franzini, et al., (2005)	Self-rated health	I: Social support C: Collective efficacy (social cohesion, and trust and helpfulness at the community and informal social control)	3,203 respondents in 100 census block group, Texas Department of Health	The effect of neighbourhood impoverishment on health is mediated by social and physical neighbourhood characteristics
Drukker et al., (2005)	Adolescents' perceived health	C: ISC (the respondents' perception of their neighbourhood) and SC&T (bonds and trust among the residents of the neighbourhoods)	801 young adolescents living in 343 Chicago neighbourhoods from the Project on Human Development in Chicago Neighbourhoods, USA and 533 young adolescents living in 36 Maastricht neighbourhoods from the Maastricht Quality of Life Study, the Netherlands	Lower socioeconomic deprivation scores and higher levels of ISC as well as SC&T were associated with higher levels of children's perceived health, in both Maastricht and the Chicago Hispanic sub-sample, but not in the Chicago non-Hispanic samples.
Veenstra (2005)	Self-rated physical health	I: Social and political trust and participation C: Social and political trust and participation	105,699 individuals in 25 communities in British Columbia, Canada,	Only a measure of depressive symptoms had variability that could be reasonably attributed to the level of the community. The social capital of communities in the form of the availability of public spaces explained some of this variability, but in the direction contrary to expectations.
Fisher et al., (2004)	Walking activity	C: social cohesion, help close-knit, trust, get along with each other, values	582 elderly residents from 56 neighbourhoods of Portland, Oregon, Collected data via computer assisted telephone interview procedures and direct mailing	Social cohesion associated with increased levels of physical activity
Merlo et al., (2003)	Use of hormone replacement therapy (HRT) and anti-hypertensive medication (AHM)	I: Social participation C: Social participation	15,456 women aged 45-73 nested within 95 neighbourhoods, Malmo Diet and Cancer Study (1991-1996)	Low individual social participation associated with lower use of HRT; but unrelated to AHM. Low neighbourhood social participation associated with lower use of HRT; but higher use of AHM
Wen et al., (2003)	Self-rated health	C: Collective efficacy (reciprocity, density of local networking social cohesion informal social control)	8,782 individuals in 343 neighbourhood clusters in Chicago, 1994 Project on Human Development in Chicago Neighbourhoods, 1991-2000 Metropolitan community Information Centre-Metro Survey	Neighbourhood social resources associated with better individual self-rated health. Presence of affluent residents sustains neighbourhood social organization, positively affecting health

3 I: Individual measures, C: Community measures.

Linden et al., (2003)	Children's metal health service use	C: Informal social control, social cohesion and trust	262 children (56 cases and 206 controls) in 36 neighbourhoods, Maastricht Mental Health Case Register	Social capital measures did not exit main effect. Strong trust and social cohesion between citizens in a neighbourhood mitigate the effect of socioeconomic deprivation on children's mental health service use.
Lindstrom et al., (2003)[a]	daily smoking	I: Social participation C: Electoral participation	5,422 individuals 77 administrative areas in Malmo, 1994 cross-sectional Malmo Public Health Survey, Sweden	Daily tobacco smoking is mainly affected by individual factors, and that the small neighbourhood differences in daily tobacco smoking in Malmö are explained by individual factors.
Lindstrom et al., (2003)[b]	Leisure time physical inactivity	I: Social participation C: Electoral participation	3,001 individuals from 68 administrative areas in Malmo, 1994 cross-sectional Malmo Public Health Survey, Sweden	Leisure time physical inactivity mainly affected by individual factors
Lindstrom et al., (2003)[c]	Self-reported sense of insecurity	I: Social participation C: Electoral participation	3,001 individuals from 68 administrative areas in Malmo, 1994 cross-sectional Malmo Public Health Survey, Sweden	Electoral participation may partly explain individual sense of insecurity related to crime in the neighbourhood.
Drukker et al., (2003)	General health, mental health, self-esteem, satisfaction and behaviour	C: Informal social control, social cohesion and trust	7,236 children from 36 neighbourhoods in Maastricht, Longitudinal cohort study of children in Maastricht's neighbourhoods	Social capital non-specifically associated with children's general health and satisfaction. Children's mental health and behaviour associated with the degree of informal social control in the neighbourhood
Hendryx et al., (2002)	Access to health care	C: General social capital (trust, efficacy, personal safety, reciprocity, voting participation and civic engagement)	19,672 individuals from 22 metropolitan statistical areas in the US, 1996 Household Survey of the Community Tracking Study	Community social capital facilities better access to health care
Browning and Cagney (2002)	Self-rated physical health	C: Collective efficacy (reciprocity, density of local networking, social cohesion, informal social control)	2,218 adult individuals in 333 neighbourhood clusters of Chicago, 1994 Project on Human Development in Chicago Neighbourhoods, 1991-2000 Metropolitan Community Information Centre-Metro Survey	Higher levels of neighbourhood collective efficacy associated with better overall health
Subramanian, Kim, Kawachi, (2002)	Self-rated health	I: Social trust C: Social trust	21,456 individuals in 40 US communities, 2000 Social Capital Benchmark Survey	Higher levels of community trust associated with better health amongst high-trust individuals. Lower levels of community trust associated with worse health for low trust individuals
Subramanian, Kawachi, Kennedy, (2001)	Self-rated health	C: (mis)trust	144,692 adults in 39 US states, 1994 Behavioural Risk Factor Surveillance Systems	Probability of reporting poor health associated with state levels of mistrust. Association found for all income groups
Sampson et al., (1997)	Violent crime and homicide	C: Collective efficacy	8,782 individuals in 343 neighbourhood clusters in Chicago, 1995 Project on Human Development in Chicago	Collective efficacy negatively associated with neighbourhood variations in violent crime and homicide

2 Wilkinson's aggregate analysis is potentially flawed for when there is a non-linear relationship between income and mortality, a potentially spurious aggregate relation will be found between mean mortality and the within-country variability (that is inequality) in income. A study is therefore needed whereby an appropriate measure of health is related at the individual level to income and simultaneously to average income and income inequality at the aggregated level. Given that Wilkinson developed his hypothesis in relation to countries as societies, it is crucial to examine a wide variety of countries with a wide range of average income and inequalities. Moreover, as the Wilkinson argument implies that once countries reach a certain stage of development (which he operationalizes as GNP $5000 per capita in 1990), inequality takes over from per capita income as the primary determinant of health, so it is important that the hypothesis link between health and income inequality is examined over time as countries become more developed. This is the subject matter of Chapter Four.

3 Happiness, life satisfaction as well as self-rated health, are among vital components of subjective states of well-being. Wilkinson considers that these three variables are the result from income inequality. Countries where there is greater inequality, experience greater unhappiness, poorer life satisfaction and consequent poorer health. In Chapter Five, they are treated as multiple outcomes for the same individual to examine the relationship with income, at both the individual and country level, further examining the Wilkinson hypothesis that income inequality results in poor life indicators on a range of variables.

4 In Chapter Six, the relationships between self-rated health, and individual and societal trust are analyzed. As such, the chapter extends consideration of the Wilkinson relative income hypothesis into the realms of social capital and social cohesion and their effect on health. As discussed in this chapter, there is considerable recent interest in the links between growing income inequality, falling social cohesion, increasing psycho-social stress and worsening health. A pertinent example for the present analysis is the study by Subramanian *et al.,* (2002) which investigated the effect of individual and group-level social trust on self-rated health. Their findings are somewhat difficult to interpret as the 40 US communities effect was quite different for individuals expressing high and low social trust. In Chapter

Six the analysis a replicated uses multilevel models to examine the relationship between individual self-rated health and social trust both at individual and country level after taking account of individual demographic and income variables.

Chapters Four, Five and Six all use multilevel models on health and life satisfaction using individual data collected on a comparable basis at the global level; the aim of this analysis is a direct test of the Wilkinson hypothesis using micro data on individuals and macro data on relative inequalities.

Chapter 3

Modelling trajectories of global life expectancy

3.1 INTRODUCTION

The overall aim of this chapter is to evaluate the Wilkinson hypothesis, at the global scale using aggregated data. But unlike previous work, the analysis is going to be a dynamic one. That is, the chapter assesses the effect that income and inequality have on the *trajectory* of life expectancy experienced by countries, during the period 1970-2002.

More specifically this chapter will:

- identify broad patterns of change in life expectancy over 30 years;
- simultaneously examine between- and within-country variation to assess the degree to which patterns of life expectancy are becoming more or less similar at national and sub-national scales;
- categorize countries into a small number of groups with distinctive trajectories, thereby identify those with improving and those with deteriorating expectancy in relation to the general global trend;
- evaluate the effects of income and of income inequality through time on life expectancy trajectories for countries during the study period.

The modelling will be done for males and females combined, although some differences between the sexes will also be commented on, based on detailed modelling not fully presented here.

The chapter is in three parts. First, I outline the data that are used in the analyses and consider how these need to be manipulated to achieve the aims outlined above. Second, the methodology that will be used to identify the life expectancy trends is described. The data have a two–level structure of occasions (annual life expectancy at birth) nested within countries. The data are first analysed using a standard, two–level, longitudinal analysis (Singer and Willett, 2003). Then, the semi-parametric random coefficient approach of Nagin (2005) is used to group and to classify countries with similar trajectories, as well as to quantify the likelihood of a particular country belonging to a specific group. The third and final part of the chapter describes a series of models, of increasing complexity, to evaluate the Wilkinson hypothesis.

3.2 DATA

This section reports on the collation, linkage and cleaning of data, as well as the methods of data estimation used to produce a comprehensive global data set. Three data sets are manipulated to provide information on the extent of health outcome, income and income inequality during the study period for 196 countries.

3.2.1 Global Life expectancy

The most reliable way to consider health variations across different nations is to focus on mortality (Shaw et al 2002, 87). In particular, we need an overall measure that summarises the mortality experience of a country. The most suitable single measure is life expectancy at birth. This is defined as the average number of years a newborn infant would be expected to live if health and living conditions at the time of its birth remained the same throughout its life. Although, in general, women have a longer life expectancy than men, they also have higher mortality rates during child-bearing years, requiring us to differentiate by gender. In practice, life expectancy at birth is a complex indicator which has to be calculated through a series of formulae from a life table. Life expectancy can be measured for different sub-populations, such as age group, gender and race. Using the United States in 1996 as an example, the life table of estimated life expectancy for each age group but with both genders combined is given in Table 3.1.

The rows of this table represent different age groups (x) for a year interval, while the columns represent different terms used in the calculation of life expectancy, so that the first row of the last column refers to the life expectancy, $e_{(x)}$, of those (babies) aged 0 to 1 based on when they were first born (x = 0). The column gives the overall life expectancy at a particular age, in this case for those aged under 1. From left to right, the columns are as follows, where x represents a specific age interval:

- the age interval concerned;
- $q_{(x)}$: proportion of the cohort dying during the age interval;
- $l_{(x)}$: number living at beginning of the year, surviving from the year before;

- $d_{(x)}$: number of people who died during the year;
- $L_{(x)}$: the number of person-years lived for a year and
- $T_{(x)}$: the total person-years that people of age x will live on under the same condition.

In the construction of the life table, it is important to realise that the same procedures are employed for all age groups in the calculation of $L_{(x)}$, except the youngest and oldest cases. The life table begins with an assumption about the population at the beginning of the year of age 0, $l_{(0)}$ - 100,000 new born babies in this case. The age-specific mortality rate, $q_{(x)}$ is now applied according to the infant mortality rate, that is $q_{(0)} = 0.00732$. This suggests that 732 ($d_{(0)}$) infants will die during their first year; hence 99,268 survive into their second year, age 1-2, giving the estimation of $l_{(1)}$. The next column, $L_{(x)}$ gives the number of person-years contribution for every year. This is made up of a whole year for those 99,268 who are still alive and a fractionally contribution for those who have died during the year. By convention, it is assumed that people die on average half way through the year, but for those under the age of 1, it is usually assumed that they live only 10 percent through the year, because most of these deaths are close to birth. Consequently, for the first age group, the number of person-years contribution is given by $0.1 \times 732 + 99268$ equals 99,365 person years for that cohort. In the last age group, that is those aged over 100 years, there is an assumption based on how long on average a person is expected to live, given that they have already survived until 100. In the USA table shown, it is assumed that this is 2.47 years. Consequently, for the last age group, $L_{(100+)}$ is given by 2.47×1504 equals 3,714 person years.

The main purpose of the function $L_{(x)}$, is to obtain $T_{(x)}$: the total number of person-years of those alive after age x. This column is obtained by accumulating across age groups, starting at the bottom of the column for the oldest age groups. Thus, $T_{(100+)}$ is 3,714 person years, $T_{(99)}$ is 3,714 + 1,812 that is equal to 5,527 person years, where 1,812 is $L_{(99)}$. For those under 1 the total person years alive is 7,617,716 ($T_{(0)}$) which is the sum of 7,518,350 ($T_{(1)}$) and 99,365 ($L_{(0)}$). In the final column of the table, life expectancy for every person who survives is given for each age interval. This is given by the total cumulative person-years alive $T_{(x)}$, divided by the total number of people who are alive, $l_{(x)}$. For instance, the expectation of life at age 0 in 1996 is $e_{(0)}$ equals 76.18 which is the total number of persons-years to live (7,617,716) after age 0 to 1 divided by the number of person alive at age 0 (100,000).

Table 3.1 Life table for the total population: United States, 1996

Age	Proportion dying during age interval $q(x)$	Number living at beginning of age interval $l(x)$	Number dying during age interval $d(x)$	Stationary population in the age interval $L(x)$	Stationary population in this and all subsequent age intervals $T(x)$	Life expectancy at beginning of age interval $e(x)$
0–1	0.007320	100,000	732	99,365	7,617,716	76 18
1–2	0.000671	99,268	67	99,240	7,518 350	75 74
2–3	0.000391	99 211	39	99,192	7,419,111	74 78
3–4	0.000324	99,173	32	99,156	7,319 919	73 81
4–5	0.000256	99,140	25	99,128	7,220,762	72 83
5–6	0.000227	99,115	22	99,104	7,121,634	71 86
6–7	0.000210	99,093	21	99,082	7,022,531	70 87
7–8	0.000197	99,072	19	99,062	6,923,448	69 88
8–9	0.000180	99,052	18	99,043	6,824,386	68 90
9–10	0.000159	99,034	16	99,027	6,725,343	67 91
10–11	0.000143	99,019	14	99,012	6,626,317	66 92
11–12	0.000148	99,005	15	98,997	6,527,305	65 93
12–13	0.000194	98,990	19	98,980	6,428,308	64 94
13–14	0.000291	98,971	29	98,956	6,329,327	63 95
14–15	0.000424	98,942	42	98,921	6,230,371	62 97
15–16	0.000675	98,900	67	98,871	6,131,450	62 00
•						
•						
•						
84–85	0.088558	36,857	3,264	35,225	244,734	6 64
85–86	0.097012	33,593	3,259	31,963	209,509	6 24
86–87	0.106082	30,334	3,218	28,725	177,546	5 85
87–88	0.115966	27,116	3,144	25,544	148 821	5 49
88–89	0.126783	23,971	3,039	22,452	123,278	5 14
89–90	0.138706	20,932	2,903	19,480	100,826	4 82
90–91	0.151603	18,029	2,733	16,662	81,346	4 51
91–92	0.165104	15,296	2,525	14,033	64,683	4 23
92–93	0.179012	12,770	2,286	11,627	50,651	3 97
93–94	0.193458	10,484	2,028	9,470	39,023	3 72
94–95	0.208669	8,456	1,764	7,574	29,553	3 49
95–96	0.224655	6,691	1,503	5,940	21,980	3 28
96–97	0.241180	5,188	1,251	4,563	16,040	3 09
97–98	0.257808	3,937	1,015	3,429	11,477	2 92
98–99	0.274304	2,922	801	2,521	8 048	2 75
99–100	0.290683	2,120	616	1,812	5,527	2 61
100+	1 00000	1,504	1,504	3,714	3,714	2.47

(Source: US Department of Health and Human services, 1999, p11)

A single source, the Health, Nutrition and Population Data Bank (http://devdata.worldbank.org/hnpstats) provides gender-specific life expectancies for some 232 countries over the period 1970-2002. In earlier years this is given as an integer figure, but more recently and for some countries, the values are given to one decimal place. To be consistent these are rounded to integer values throughout the analyses which follow.

The data have a 'panel-like' structure (Figure 3.1) with data for years nested within countries. However there is considerable imbalance as not all countries report every year (e.g. South Korea, where reporting intervals have changed over time) and because 'new' countries have come into existence in recent years (e.g. Macedonia, FYR became independent from Yugoslavia in 1991). Excluding any country with less than 5 measurements over the study period leaves data for 196 countries. The graph in Figure 3.2 is a plot of the raw data with a line drawn for each country for each time point. There are clearly substantial differences between countries, and there is obvious volatility over time as the lines criss-cross over each other. The data show considerable variation and this should allow for a good test of the Wilkinson hypothesis.

3.2.2 Income and inequality of countries

The life expectancy at birth data were linked with data on GDP and also data on income inequality, in the form of estimated household-income inequality coefficients from the University of Texas Inequality project (http://utip.gov.utexas.edu). These are described below, in Section 3.2.2.2

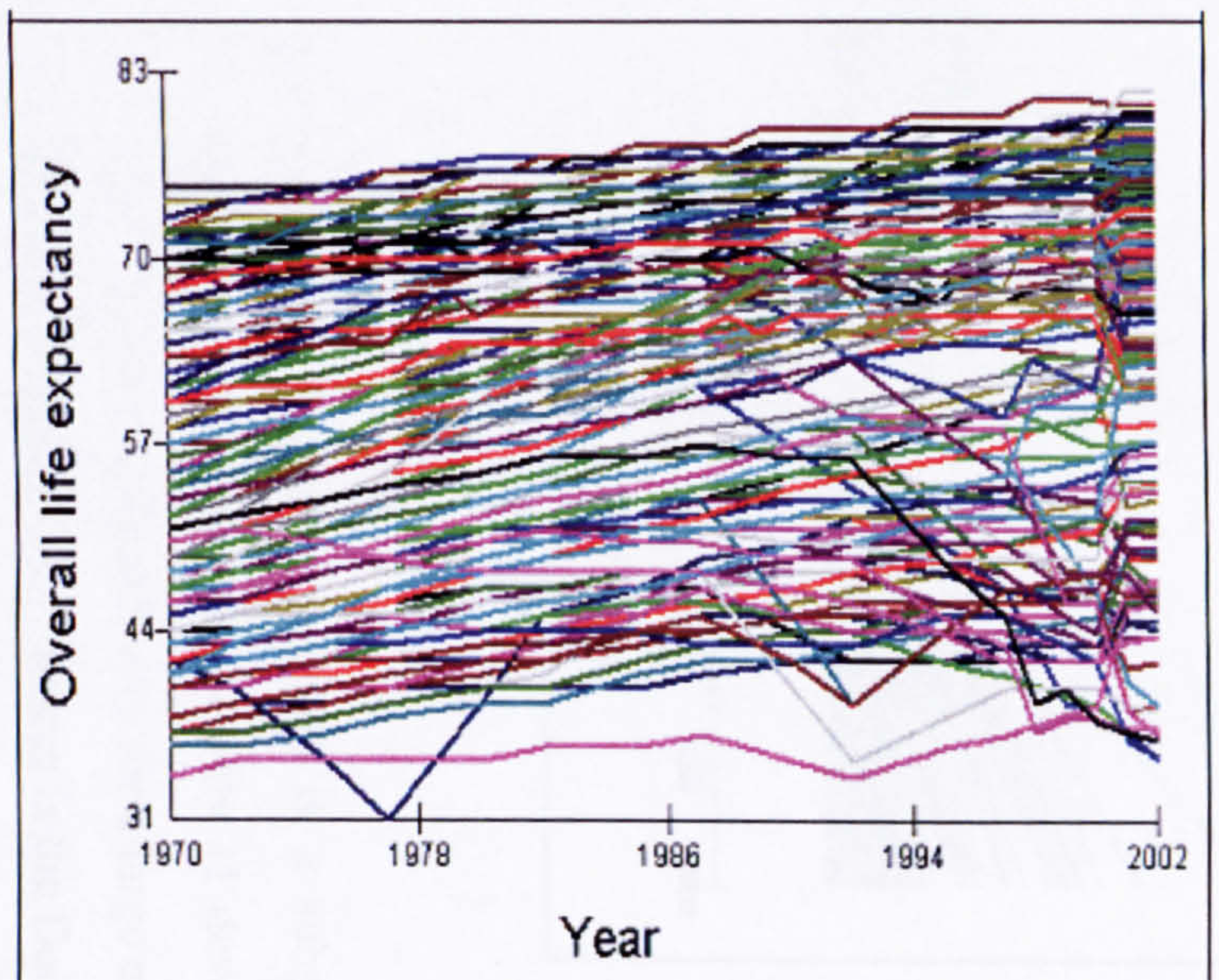
3.2.2.1 Changing GDP

The World Bank website (http://devdata.worldbank.org/hnpstats) provides international country-level income data on Gross Domestic Product per capita (GDP pc) (in purchase power parity (PPP) for 2003 US dollars) on an annual basis from 1970 to 2000 for some 232 countries, including all 196 countries used in this study. I have also added comparable data for 2001 and 2002 from the World Factbook (http://www.cia.gov/cia/publications/factbook). As stated in Chapter 2, according to the definition from the *Human Development Report 2001* (United Nations, 2001), GDP pc represents the final product of total goods and service at market value for an economy shared by its population, including the residents and foreign labour within the territory for a period of time. PPP is a rate of exchange that accounts for price differences across countries allowing international comparisons at the PPP US$ rate. A PPP US$1 has the same purchasing power in the domestic economy as $1 has in the United States. This is clearly an ideal source for country year-specific data on a country's absolute income on a comparable basis (Firebaugh, 2003).

Figure 3.1 Panel-like (longitudinal) multilevel data structure with occasion nested within country

*71 stand for the year; 1971 of life expectancy at birth at level 1 and vice versa.

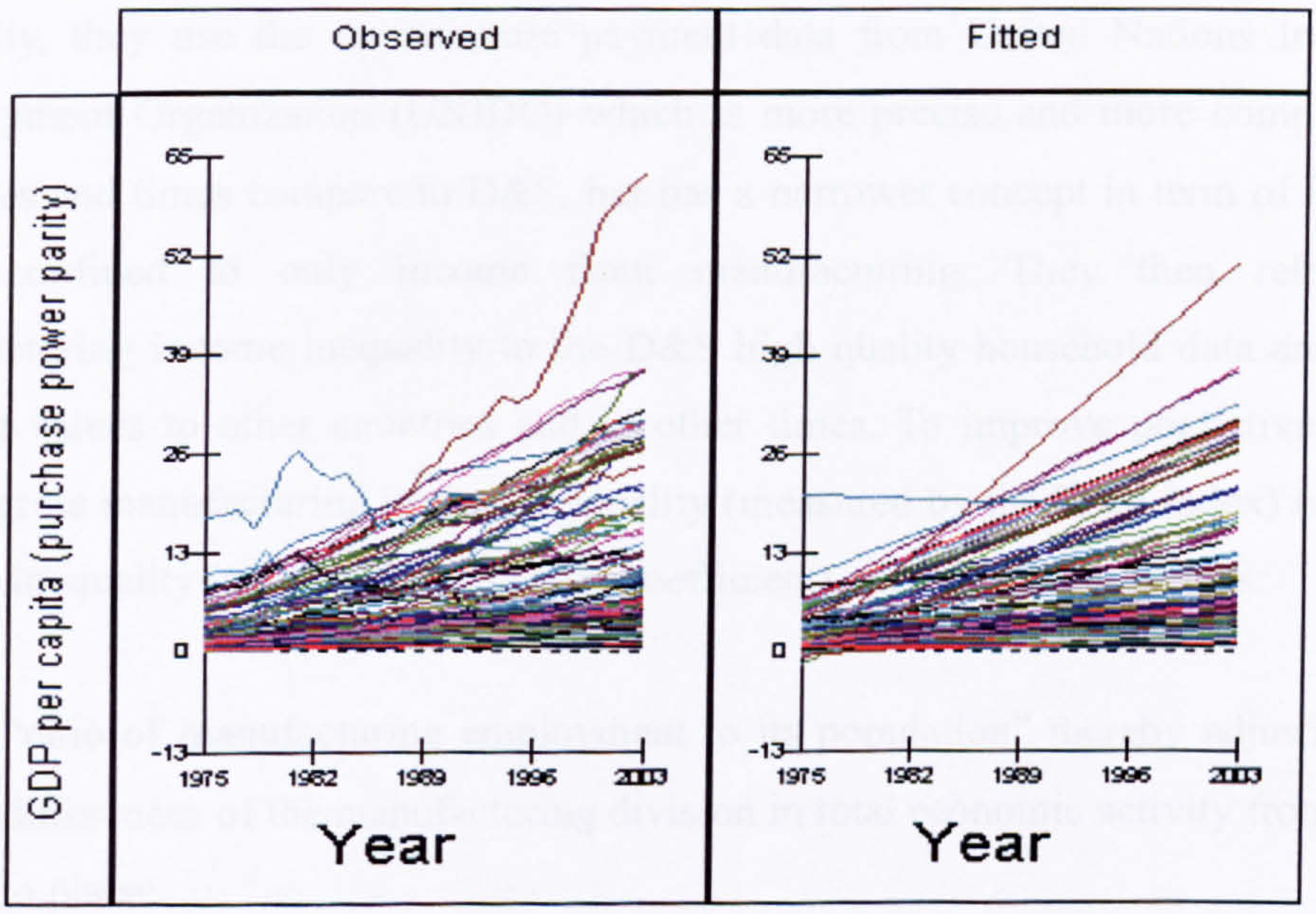
Figure 3.2 Plot of the raw data on life expectancy for both sexes combined 1970-2002



Detailed data is also required for all our countries to match our mortality data, but even with United Nations data it was found that considerable amounts were not available; therefore the following imputation strategy was adopted.

Observed GDP was regressed on a cubic polynomial of time in a two-level multilevel model (Hedeker, 2005) where level 1 is occasion and level 2 is country. This model estimated the general trend across all countries and allowed the trend to vary from place to place. Details of this type of modelling will be considered in detail later, but it is important to stress that the resultant estimates are 'precision-weighted' and 'borrow strength': that is, where there is little information on a country's specific trend it will be 'shrunk' back towards the global tend (Jones and Bullen, 1994). The predicted model estimates of the trend in each country had a correlation in excess of 0.99 with the observed trend. Figure 3.3 shows the observed and fitted trend in GDP.

Figure 3.3 The changing pattern of each country's GDP per capita (ppp) for current international dollars of observed and fitted dataset



3.2.2.2 Changing Inequality

Obtaining reliable comparable information on income inequality is a substantial challenge. Wilkinson, in his work often relies on information for only 11 developed countries. We require inequality to be measured in a consistent way for a large number of countries over time to answer our research questions. A key source is the Deininger and Squire (D&S) information provided by the World Bank (http://www.worldbank.org/research/growth/absineq.htm, 1996), which has often been used as the major source of data in previous studies of income inequality, but not of health outcome. This data have been graded (by Deininger and Squire, 1996) into different levels of reliability and there are some 700 country/year observations in the "high quality" version data set. Unfortunately these data are highly selective with just a few countries providing annual or nearly annual observations over long periods of time. The data are not only deficient in coverage but also, even for the high-quality data, inequality has been measured inconsistently; some are net income values, others are based on expenditure; some are per household and some are per person. Consequently, international comparison cannot be done with the D&S dataset because there are too many gaps and too much error in their measurements.

An attractive alternative source of income inequality data comes from the UTIP-UNIDO project at the University of Texas (http://utip.gov.utexas.edu/web/atip.htm). This has some 3,200 observations of income inequality over 36 years (1963-1999). Basically, they use the manufacture payment data from United Nations Industrial Development Organization (UNIDO) which is more precise and more complete for countries and times compare to D&S, but has a narrower concept in term of income, being confined to only income from manufacturing. They then relate the manufacturing income inequality to the D&S high quality household data and input missing values to other countries and to other times. To improve predictive power they regress manufacturing income inequality (measured by the Theil index) on D&S income inequality (measured by the Gini coefficient) and three other factors:

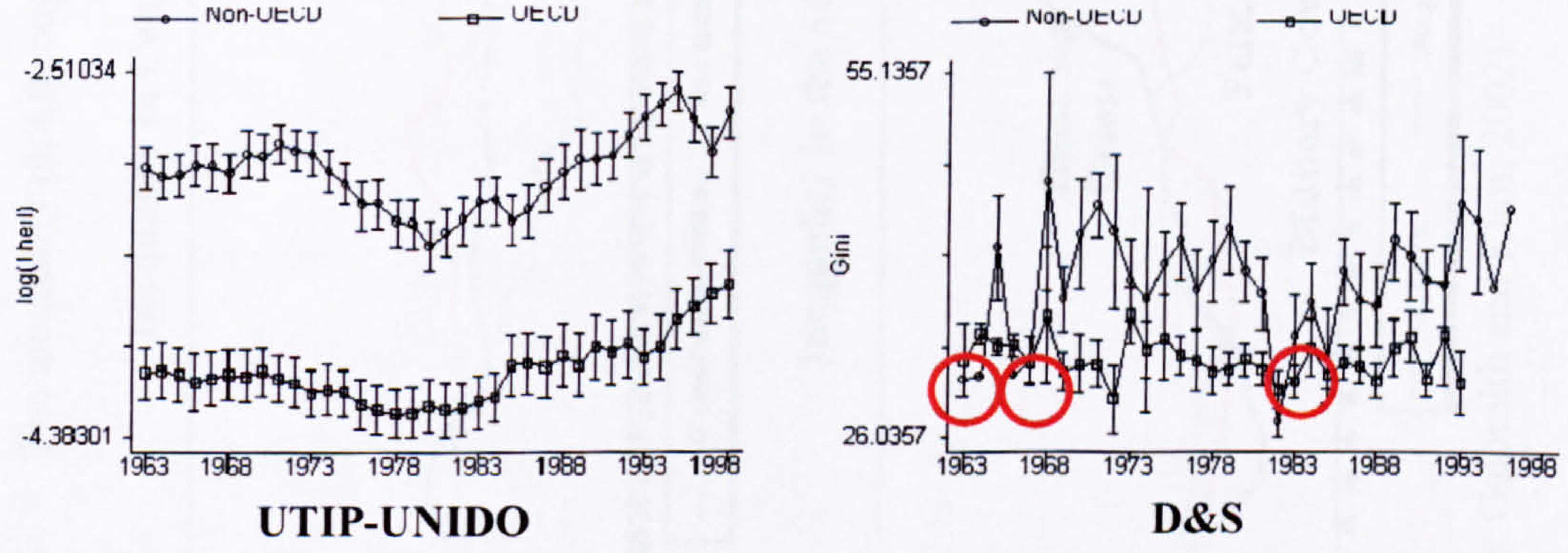
- "ratio of manufacturing employment to its population" thereby adjusting the differences of the manufacturing division in total economic activity from place to place;
- "the share of urban population" wealthy people live in cities hence urbanization recognizing rising inequality of income;
- "population growth rate", to reflect the age structure

.There are also three dummy variables representing four different types of income data in the D&S:

- household gross income (as base),
- expenditure income;
- per capita income;
- net income.

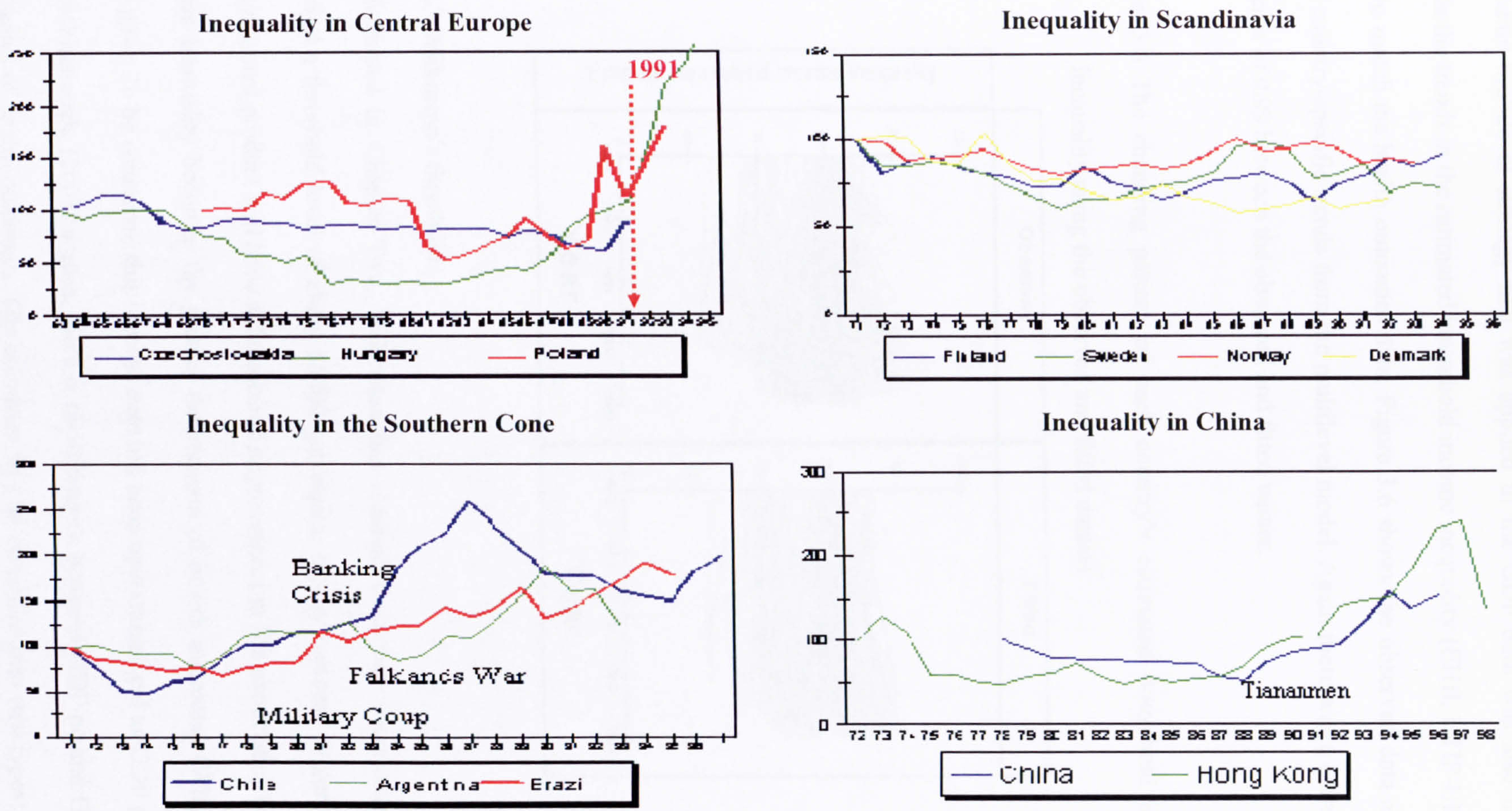
The resultant and well-fitting model in diagnostic terms (it has a R-square of 0.63) is then used to predict the Gini inequality on the basis of the predictor variables, as the dispersion of manufacturing pay is available for a large number of countries. Galbraith and Kum (2004) undertake a number of qualitative comparisons to confirm that their predictions are an improvement on the lower-quality D&S data. For example, average inequality should be higher in Non-OECD[1] rather than OECD countries, but this is not the case in the three years (circled in red) using D&S data (see Figure 3.4). It is also apparent that the D&S data show a very high and unlikely year-on-year variation. In contrast, the UTIP-UNIDO data show patterns that are expected over time according to historical development and specific historical events. Figure 3.5 shows a number of examples: the increasing inequality in the former Soviet Union countries in central Europe; increases following the Tiananmen Square incident in China; and worsening inequality in Southern America countries associated with the banking crisis, Falklands War and a military coup. Scandinavia has the expected stable distribution of income across the years under its well-established, social-welfare system.

Figure 3.4 Comparison of the average trends of inequality in UTIP-UNIDO and D&S data between Non-OECD and OECD countries



[1] OECD: The groups of 30 member countries sharing a commitment to democratic government and the market economy which have active relationships with some 70 other countries, NGOs and civil society.

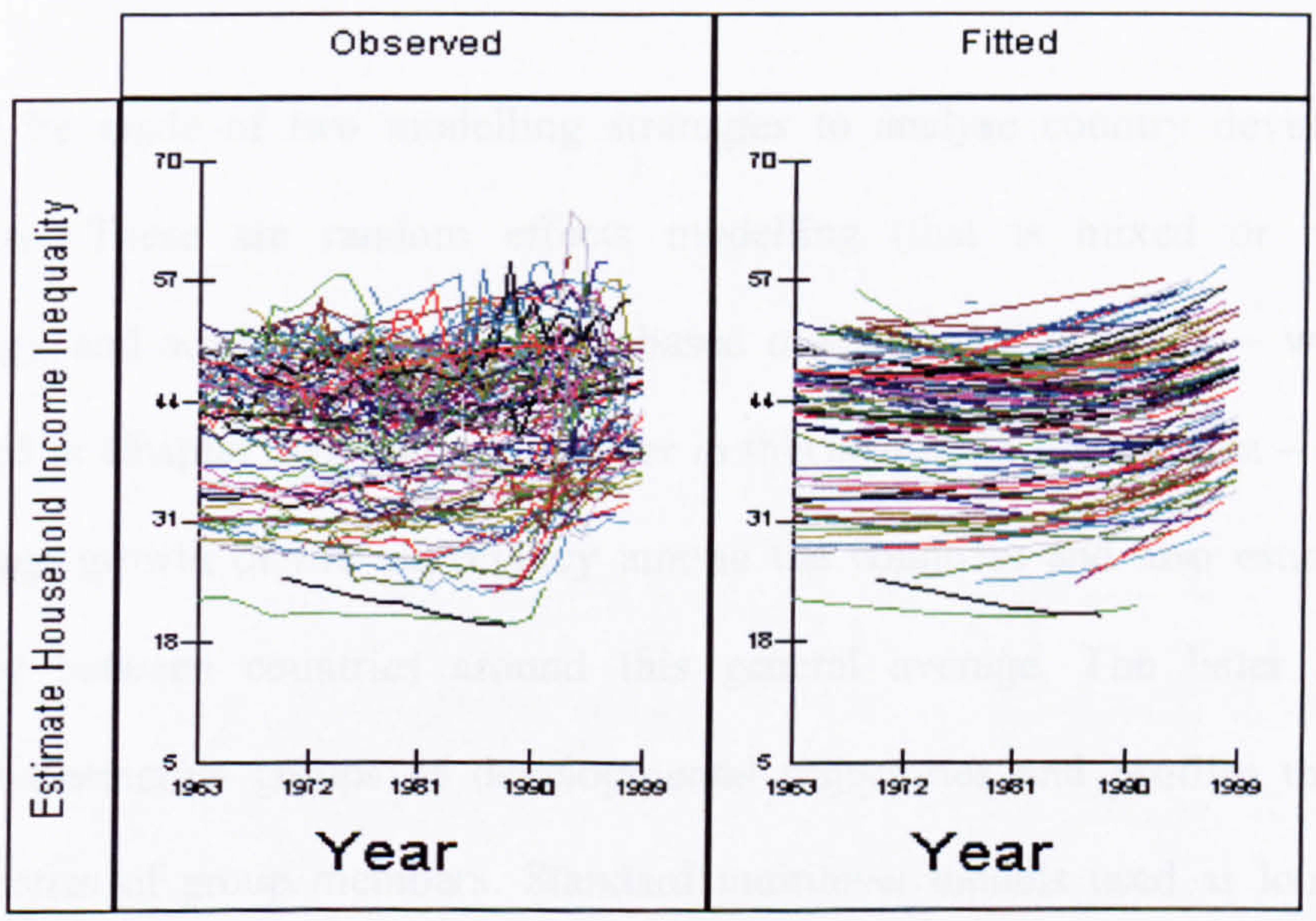
Figure 3.5 The inequality plot through time for a number of countries in UTIP-UNIDO



(Resource: Galbraith and Kum, 2002)

The same multilevel strategy that was applied to the GDP data was also used to impute the trends in the estimated household income inequality (EHII, UTIP-UNIDO) data to match the health outcomes data. Figure 3.6 shows the observed data and the fitted country-specific trends from the multilevel model. Again there was a correlation in excess of 0.99 between the observed and fitted values.

Figure 3.6 The changing pattern of each country's estimated household income inequality using the observed and fitted dataset



3.2.2.3 Wilkinson's threshold

As discussed in Chapter Two, Wilkinson has claimed that once economies have achieved a threshold level of about \$5000 per capita, further increase in per person gross national product will have a diminished improvement in life expectancy because income inequality becomes the greater determinant of health outcomes (Wilkinson, 1996 p34). To be consistent this income data has been operationalised as GDP pc PPP for, as Firebaugh, (2003) argues, there is no difference between GDP pc and GNP pc for studies of health outcomes. The countries will be classified into two types: above and below this threshold according to their GDP pc PPP in 1990 which is the same year that Wilkinson used in his 1996 book.

3.3 METHODOLOGY

In this section, I discuss an appropriate methodology for

- identifying the trajectories in mortality over time for the 196 countries;
- relating those trajectories to GDP of and income inequality within the countries, thereby evaluating the Wilkinson hypothesis.

Use will be made of two modelling strategies to analyse country developmental trajectories. These are random effects modelling (that is mixed or multilevel modelling), and semi-parametric, group-based modelling. The former – which was introduced in Chapter Two and used earlier in this chapter to impute data – estimates the average growth of life expectancy among the countries and also estimates the variability between countries around this general average. The latter approach identifies distinctive groups of developmental trajectories and profiles the shared characteristics of group members. Standard multilevel models used as longitudinal models have now received extensive treatment in the literature (e.g. the book-length treatment of Singer and Willet 2003). However, the group approach is a more recent development and will therefore be given a more extended treatment here.

3.3.1 Random effect modelling

3.3.1.1 Overall Life expectancy

Clearly there is a hierarchical structure to the dataset collected for this chapter, in that life expectancy measurements through the study period can be conceived as being nested within each country. There are also imbalances ('missing data') that will have to be taken into account. Multilevel modelling can deal with such imbalances and hierarchical structure. It can also deal with the need to model both country and occasions simultaneously, to get an adequate description of the trends (i.e. country level differences over and above global trends). A standard multilevel model for change, a simple model with two levels, has been specified as follows, to identify broad patterns of change in life expectancy:

Response	Fixed	Random
$Y_{ij} =$	$\beta_0 X_{0ij} + \beta_1 T_{ij} + \beta_2 T_{ij}^2 + \beta_3 T_{ij}^3 +$	$(\mu_{0j} X_0 + \mu_{1j} T_{ij} + \varepsilon_{0ij} X_0 + \varepsilon_{1ij} T_{ij})$

The dependent variable Y_{ij}, is overall life expectancy (LE) for year i in country j and is related in the fixed part of the model to a cubic polynomial of time.[2] The variable T, which represents the orthogonal polynomial of the calendar year, has been centred around 1986 so that β_0 estimates the global average life expectancy in 1986 and β_1 is the global linear increase in LE per annum. The use of polynomials allows a non-linear trend over time; a cubic polynomial allows up to two changes of direction; orthogonal polynomials avoid multi-collinearity problems, which would occur in such models if only the square and cube of time were used (Hedeker and Gibbons, 2005). Moreover, the use of orthogonal polynomials means that each predictor is on the same scale which allows a direct comparison of the magnitudes of the estimated coefficients.

In the random part of the model, there are terms involving linear time at both levels.[3] The μ_{0j} term gives the difference for each country from the global mean in 1986, the

---

[2] More complex functions of time were evaluated but were found not to be needed in the fixed part.

[3] More complex functions of time were evaluated but were found not to be needed in the random part.

so-called random intercept. If μ_{0j} is positive, then a country has a longer life expectancy than the global mean (β_0) in 1986; if negative, then it has a worse mortality. The μ_{1j} gives the differential trajectory, or rate of change over time for each country; the so-called random slope. If positive it means that the country's life expectancy is improving more steeply than the global trend, but a negative value suggests a general worsening in relation to the global trajectory.

These random intercept and slope terms at level 2 are assumed to come from a joint multivariate normal distribution for countries

$$(\mu_{0j}, \mu_{1j}) \sim \mathrm{N}(0, \Omega_\mu), \qquad \Omega_\mu = \begin{bmatrix} \sigma^2_{\mu 0} & \\ \sigma_{\mu 01} & \sigma^2_{\mu 1} \end{bmatrix}$$

This consequently implies that the between-country variance around the global average is given by a quadratic function in time:

$$\sigma^2_{u0} + 2\sigma_{u0\,u1} T_{ij} + \sigma^2_{u1} T_{ij}^2$$

Such a quadratic variance function allows for between-country differences to increase, decrease or indeed remain unchanged through time. The covariance term in such models can be given a direct interpretation: if positive then countries with high life expectancy have better improving trajectories and cause even more dissimilarity between countries; if negative, it means countries with higher life expectancy are losing their relative position over time and so countries will converge toward greater homogeneity.

The level 1 random terms are ε_{0ij}, which gives the allowed to vary departure from the country specific estimated life expectancy in 1986, and ε_{1ij} which gives the departure from the country trend on each occasion. Again it is assumed that the occasion specific random terms come from a joint multivariate normal distribution:

$$(e_{0j}, e_{1j}) \sim N(0, \Omega_e), \qquad \Omega_e = \begin{bmatrix} \sigma^2_{e0} & \\ \sigma_{e01} & \sigma^2_{e1} \end{bmatrix}$$

This implies that the within country variation is also given by a quadratic function of time:

$$\sigma^2_{e0} + 2\sigma_{e0\,e1}T_{ij} + \sigma^2_{e1}T_{ij}^2$$

Such a quadratic variance function allows for between occasion within country differences to increase, decrease or indeed remain unchanged through time. All the results of this model have been obtained using the software MLWiN (Rasbash *et al.*, 2000).

The results from applying the two-level model of overall life expectancy at birth of 196 countries from year 1970 to 2002 are given in Table 3.2. It can be seen in the fixed part that the global mean life expectancy in 1986 (β_0) was 63.05 years, and that a cubic polynomial is needed to describe the general trend, as all the z-ratios for the fixed part coefficients exceed the 0.05 cut-off of 2. A linear increase in β_1 of 0.25 per year means that in every 4 years over the period, people live on average 1 year longer.

In the random part, there are significant effects for all terms, with the exception of the covariance term at level 2. The level 2, between country variations, suggest that there is a significant variation between countries in 1986 (σ^2_{u0} is estimated to be 109.9), and that there is a significant variation in the trajectories of countries around the global trend (σ^2_{u1} is estimated to be 0.044). The non-significance of the co-variance term ($\sigma_{u0\mu1}$ of –0.001), suggests that there is little correlation (-0.001) between a country's position relative to the global mean in 1986 and its relative trajectory around the global trend over the study period. Knowing a country's relative position in the middle of the sequence, tell us little about the relative trajectory or rate of change over the entire period.

Table 3.2 Parameter estimates obtained for the model of overall life expectancy: fixed and random parts

Fixed

Term	Estimate	SE	Z ratio	P value
β_0 (1986)	63.05	0.758	83.18	0.00
$\beta_1 T$	0.246	0.016	15.38	0.00
$\beta_2 T^2$	-3.783	0.446	8.48	0.00
$\beta_3 T^3$	-0.834	0.232	3.59	0.00

Random

Level	Terms	Estimates	SE	Correlation	Chi-square	P value
2	σ^2_{u0}	109.9	8.549	1	165.222	0.00
2	$2\sigma_{u0\,u1}T$	-0.001	0.159	-0.001	0.000	1.00
2	$\sigma^2_{u1}T^2$	0.044	0.005	1	70.915	0.00
1	σ^2_{e0}	0.718	0.111	*	42.153	0.00
1	$2\sigma_{e0\,e1}T$	0.085	0.111	*	28.099	0.00
1	$\sigma^2_{e1}T^2$	0.017	0.003	*	34.407	0.00

These results are more easily appreciated graphically in Figure 3.7. These figures show some interesting and important results. The graph in the top left-hand plot shows the estimated global trend derived from the cubic polynomial terms of the fixed part of the model. On average, life expectancy has improved rapidly but appears to have peaked in the last decade of the twentieth century. The panel on the top right is the between-country variance which shows a marked U-shape. There used to be substantial differences between countries in the early 1970s; these converged somewhat to the global mean by the mid 1980s but since then there has been increasing differences between countries around the generally rising mean. These results are not an artefact of centring on 1986.

The bottom left panel shows the within-country variance as a J-shaped curve over time. These within- country differences are much smaller than the between-country differences (shown by the different scales on the vertical axes), and suggest that the last 15 years have been characterised by an increasing volatility in a country's trajectory over time. It is difficult to interpret this, but it may reflect increasing sensitivity of the data collection process over time.

Figure 3.7 Global trend and variance functions: overall life expectancy

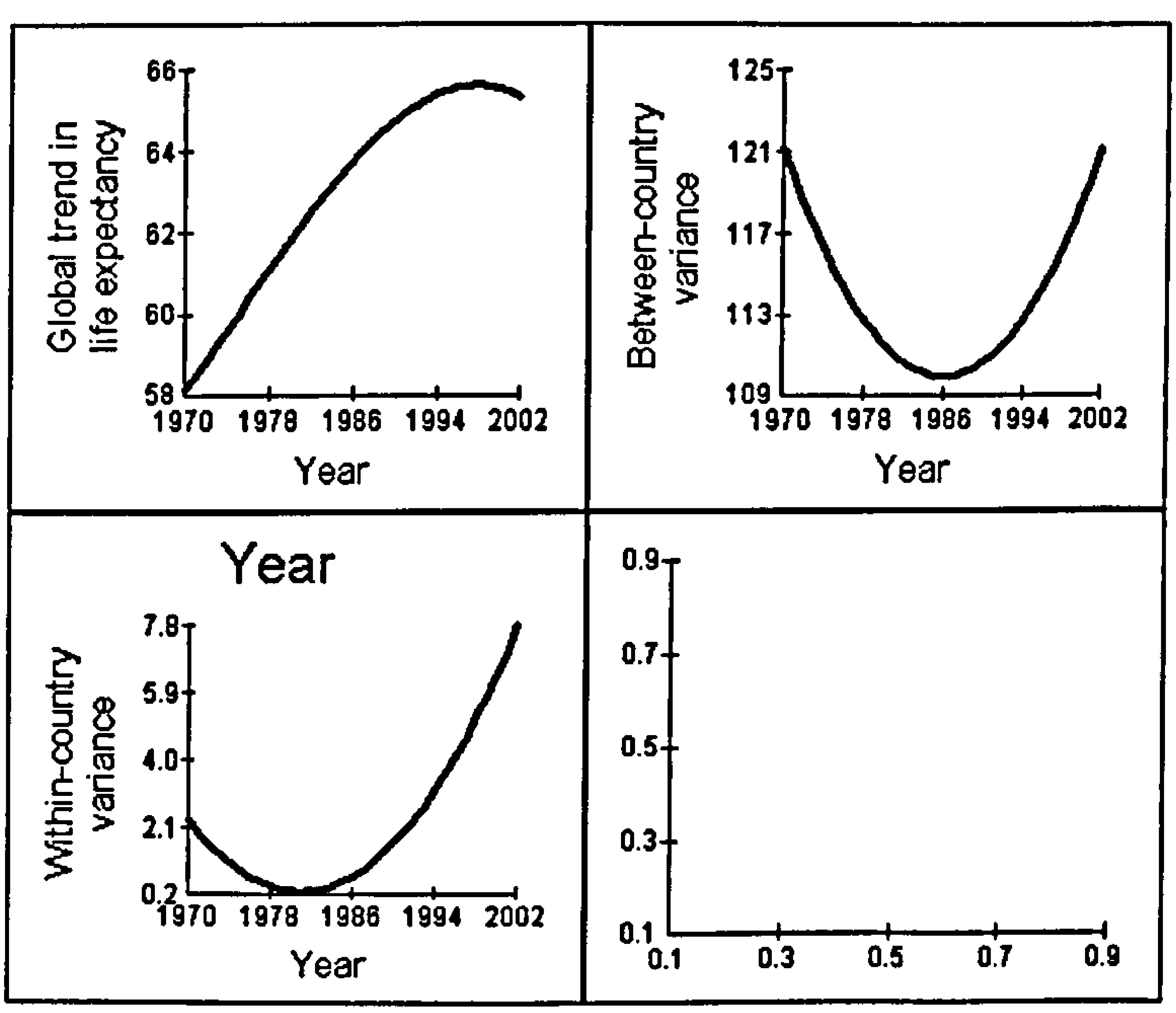


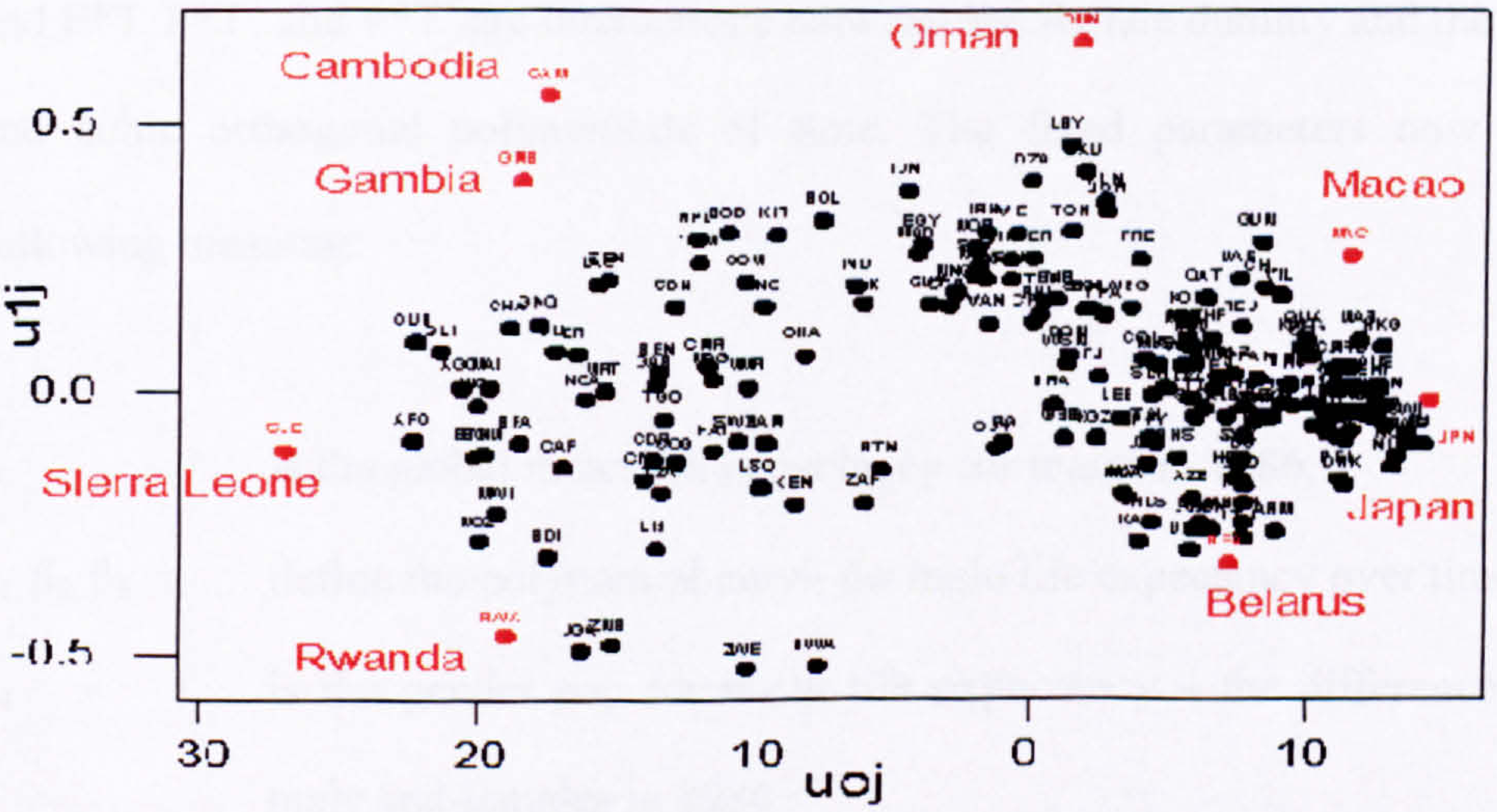
Table 3.3 gives the estimated variances for three time points at each level. The relevant importance of the between country variance in comparison to the within-country, between occasion variance is very clear.

Table 3.3 Level 1 and 2 variances at the beginning, middle and end of the time series

Between-country variances	Year	Overall
	1970	121.28
	1986	109.89
	2002	121.18
Within-country variances	**Year**	**Overall**
	1970	2.37
	1986	0.72
	2002	7.83

The individual country-level terms are plotted in Figure 3.8. This is a scatter-plot of μ_{0j} and μ_{1j} with abbreviated country names. Two countries representing the extremes of four distinctive patterns have been picked up and highlighted in red. For instance, Macao has a higher life expectancy of about 11 years in 1986 than the global average and is doing very well through time.

Figure 3.8 Scatter-plot of countries' differentials (μ_{0j} vs μ_{1j})



Oman has a similar pattern and improves tremendously. In contrast, Rwanda and Sierra Leone's life expectancy are not only less than the global average by 18 and 27 years respectively but they are also getting worse. Japan had the highest longevity in 1986 and over the period its trend mirrors the global average. Belarus had long life expectancy in 1986 but has a worsening experience over time. In contrast, Cambodia as well as Gambia have made very good progress from the much lower global average LE.

3.3.1.2 Life expectancy for males and females

The life expectancy data have so far been for both men and women combined. Results are now reported for each sex modelled separately as part of an overall model:

$$Y_{ij} = \beta_0 X_0 + \beta_1 T + \beta_2 T^2 + \beta_3 T^3 + \beta_4 F + \beta_5 F{*}T + \beta_6 F{*}T^2 + \beta_7 F{*}T^3 + (\mu_{0j} X_{0ij} + \mu_{1j} T + \mu_{4j} F + \mu_{5j} F{*}T + \varepsilon_{0ij} X_0 + \varepsilon_{1ij} T + \varepsilon_{4ij} F + \varepsilon_{5ij} F{*}T)$$

where Y_{ij} is life expectancy for country j at time point i; X_{0ij} is the constant; T_{ij} is time centred around 1986; T^2 and T^3 provide the cubic polynomial of time, again centred around 1986; F is am indicator variable identifying females with a 1, males with a 0; and $F{*}T$, $F{*}T^2$ and $F{*}T^3$ are interactions between the female dummy and the quadratic and cubic orthogonal polynomials of time. The fixed parameters now have the following meaning:

β_0	is the global mean life expectancy for males in 1986;
$\beta_1, \beta_2, \beta_3$	define the polynomial curve for male life expectancy over time;
β_4	is the gender gap for mean life expectancy – the difference between male and females in 1986

$\beta_5, \beta_6, \beta_7$ define the differential trend for females in contrast to males.

In the random part, there is a complex expression for the variability around the global trends between countries that is differentiated for males and females and over time

where

μ_{0j} is the country differential for males in 1986;

μ_{1j} is the linear trend for males for each country;

μ_{4j} is the differential life expectancy from males for each country for females in 1986;

μ_{5j} is the extent to which the linear trend for females in a country differs from the male trend;

and the overall variance function is

$$\sigma^2_{u0} + 2\sigma_{u0\,u1}T + \sigma^2_{u1}T^2 + 2\sigma_{u0\,u4}F + 2\sigma_{u1\,u4}T\times F + \sigma^2_{u4}F^2 + 2\sigma_{u0\,u5}F.T + 2\sigma_{u1\,u5}T\times F.T + 2\sigma_{u4\,u5}F\times F.T + \sigma^2_{u5}F.T^2$$

It was not possible to fit a more complex model in which the polynomials of time and interactions with the female contrast were included; convergence could not be achieved[4]. Table 3.4 gives the model estimates: it can be seen that, in 1986 and on average across the globe, females lived some 4.7 years longer than males. It can also be seen in the fixed part of the model that there are significant interactions with time

[4] MLwiN using an iterative fitting procedure for which convergence cannot necessarily be guaranteed for all models.

and time^2 for females. These results can best be appreciated graphically and Figure 3.9 shows the overall global trend for each sex. It would appear that after a long-term increasing of both sexes have seen a 'flattening out' of life expectancy during the 1990's that is more marked for men than women, so that the overall gender gap has increased.

Table 3.4 Model estimates: men and women separately

Fixed

Term	Estimate	SE	Z ratio	P value
β_0 (1986, Male)	60.86	0.709	85.84	0.000
$\beta_1 T$	0.243	0.015	16.20	0.000
$\beta_2 T^2$	-3.477	0.397	8.76	0.000
$\beta_3 T^3$	-1.048	0.230	4.56	0.000
$\beta_4 F$	4.682	0.157	29.82	0.000
$\beta_5 F*T$	0.019	0.005	3.8	0.000
$\beta_6 F*T^2$	-0.464	0.206	2.25	0.024
$\beta_7 F*T^3$	0.337	0.194	1.74	0.082

Random

Level	Terms	Estimates	SE	Correlation	Chi-square	P value
2	Σ^2_{u0}	96.68	7.784	1.000	154.291	0.000
2	$2\sigma_{u0u1}T$	-0.016	0.134	-0.008	0.014	0.906
2	$\Sigma^2_{u1}T^2$	0.040	0.004	1.000	89.938	0.000
2	$2\sigma_{u0u4}F$	11.44	1.132	0.542	102.044	0.000
2	$2\sigma_{u1u4}F*T$	-0.164	0.037	-0.382	20.013	0.000
2	$\Sigma^2_{u4}F^2$	4.614	0.528	1.000	76.406	0.000
2	$2\sigma_{u0u5}F.T$	0.105	0.039	0.538	7.253	0.007
2	$2\sigma_{u1u5}F.T*T$	0.003	0.001	0.705	7.132	0.008
2	$2\sigma_{u4u5}F.T*F$	-0.002	0.012	-0.045	0.027	0.869
2	$\sigma^2_{u5}F^2 T^2$	0.000	0.000	1.000	0.500	0.480
1	$\sigma^2 e_0$	0.758	0.108		49.513	0.000
1	$2\sigma e_0 e_1 T$	0.064	0.012		29.498	0.000
1	$\sigma^2 e_1 T^2$	0.014	0.002		44.665	0.000
1	$2\sigma e_0 e_4 F$	-0.022	0.033		0.434	0.510
1	$2\sigma e_1 e_4 F*T$	0.033	0.009		13.496	0.000
1	$\sigma^2 e_4 F^2$	0	0		0	*
1	$2\sigma e_0 e_5 F.T$	0	0		0	*
1	$2\sigma e_1 e_5 F.T*T$	0.003	0.001		15.083	0.000
1	$2\sigma e_4 e_5 F.T*F$	0	0		0	*
1	$\sigma^2 e_5 F^2$	0	0		0	*

The random parts of the models are again best appreciated graphically. The right top panel and the left bottom panel in Figure 3.8 show the level-2 variance of between-country and level-1 variance of between-year, respectively. There are basically the same for both sexes and have seen an increase since the early 1980s. Consequently, for both males and females, within-country life expectancy has become more volatile over time. There is a notable between countries differences with females consistently experiencing greater variation. Countries matter a great deal for male life expectancy, but they matter even more for females. There general U shaped pattern is, however, found for both sexes with increasing divergence in rates around the global average being experienced by both sexes since the middle 1980s. The between and within country variances are given in Table 3.5. On the basis of combining these differentials and the fixed estimates, Figure 3.10 plots the estimated trajectories for all 196 countries over the period.

Figure 3.9 Looking at life expectancy for male and females: the Global trend and variance functions

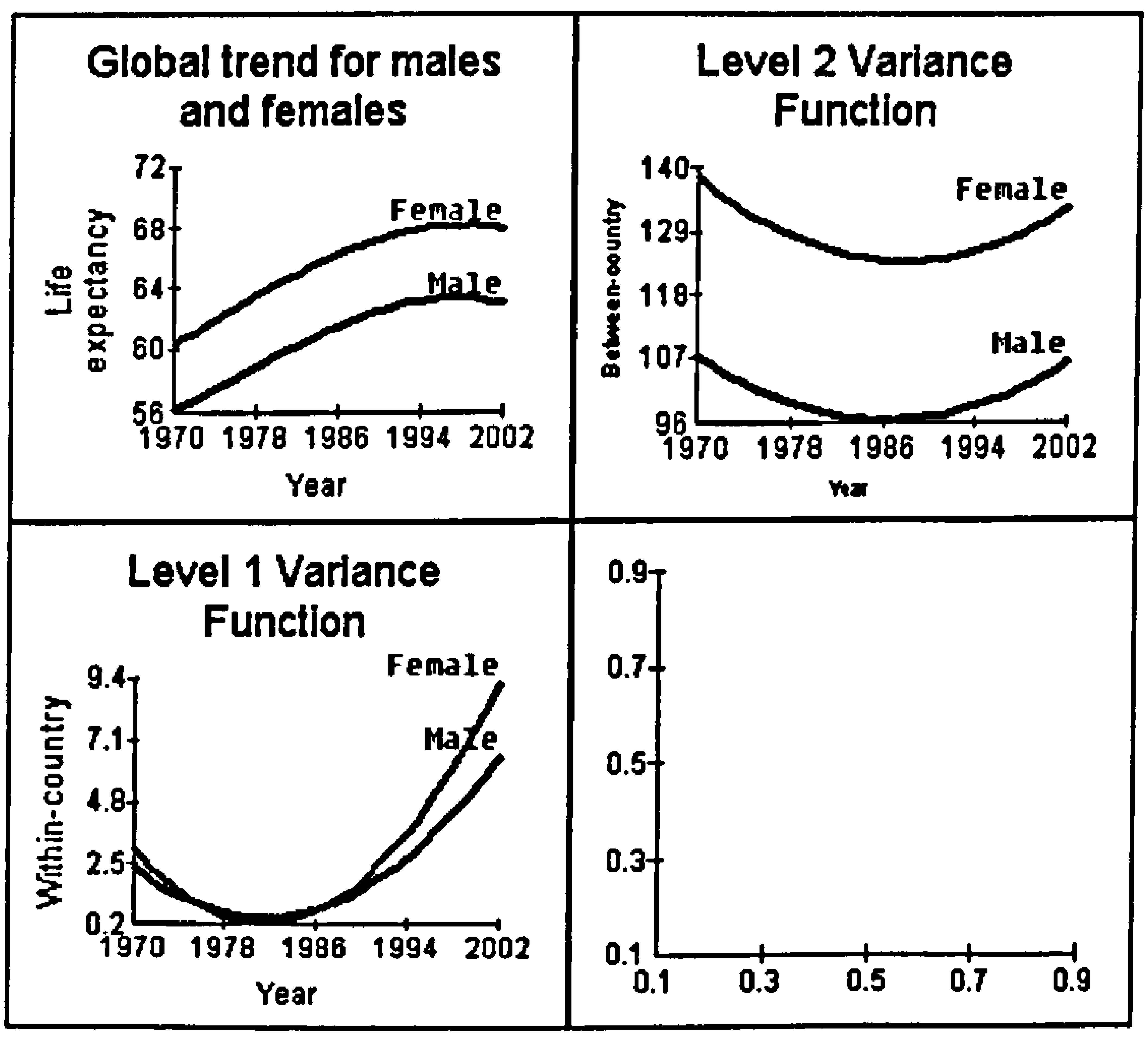


Table 3.6 lists the extreme differentials for the μ_{0j}, μ_{1j}, μ_{2j} and μ_{3j} terms. Thus, in comparison to a global average of 61.5 years for men in 1986, Japanese men live 13.81 years longer, while those in Sierra Leone live 25.79 years shorter. Globally women live 4.68 years more than men, but in Russia there is an extra differential of 5.80 years, while in Nepal the gender gap is reversed (men live longer than females).

Table 3.5 The Fixed estimates: between country and within country variances for 1979, 1986 and 2002

Between country variances	Year	Male	Female	Gender gap
	1970	107.375	138.327	30.952
	1986	96.683	124.171	27.488
	2002	106.343	133.440	27.097
Within country variances	Year	Male	Female	Gender gap
	1970	2.341	2.999	0.658
	1986	0.758	0.714	0.044
	2002	6.452	9.223	2.771
Fixed estimate	Year	Male	Female	Gender gap
	1970	56.138	60.228	4.090
	1986	61.534	66.306	4.772
	2002	63.105	68.049	4.944

Figure 3.10 Country trend for male and female life expectancy

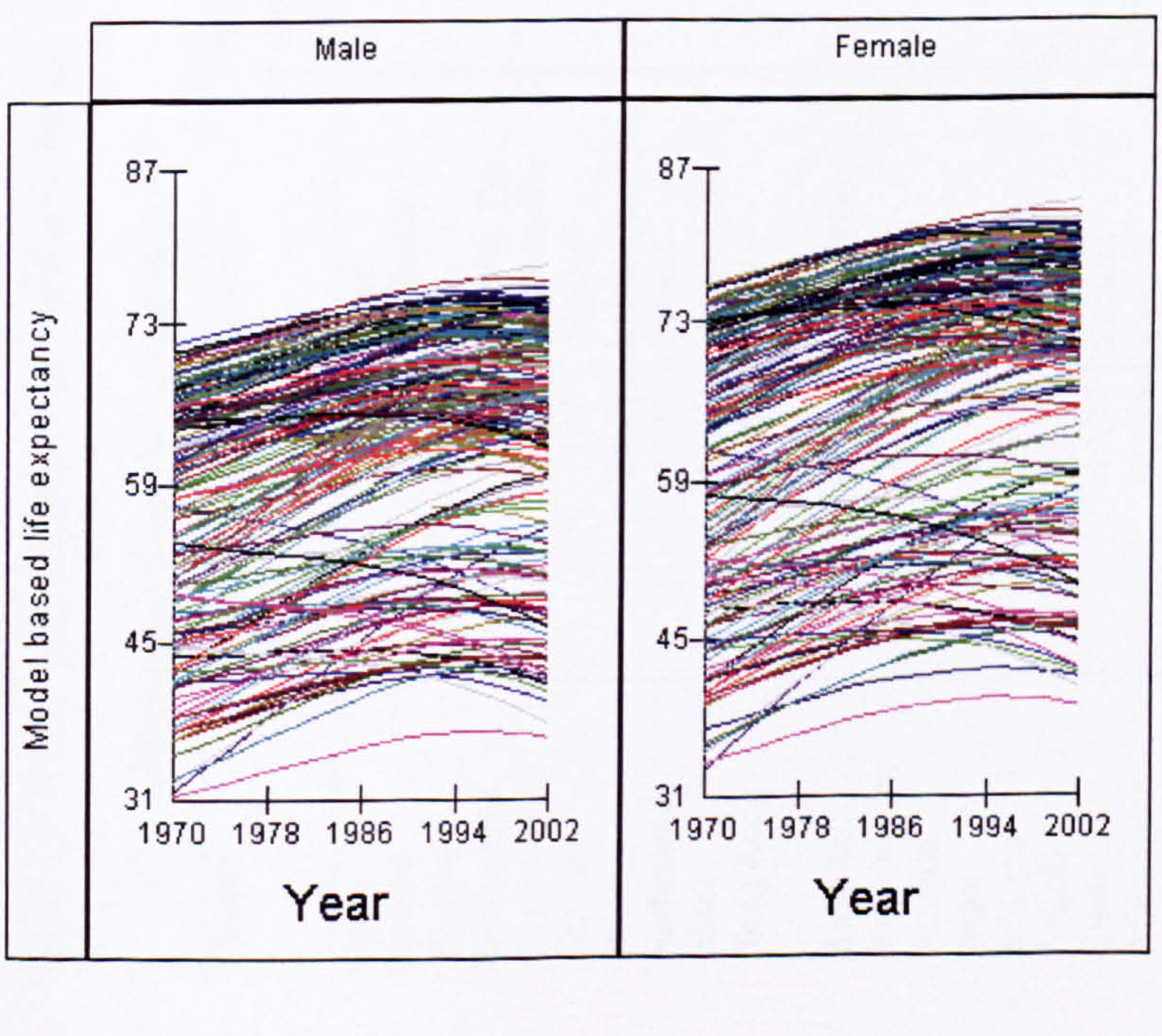


Table 3.6 Country differentials in male and female life expectancy: the ten worst and ten best countries in terms of difference in 1986 and differential from the global trend

Country	Male differential from global average in 1986	Country	Male differential from global linear trend 1970-2002	Country	Female differential from global average in 1986	Country	Female differential from global linear trend 1970-2002
Japan	13.74	Oman	0.60	Russia	5.78	Oman	0.039
Iceland	13.62	Cambodia	0.51	Latvia	4.94	Tunisia	0.029
Sweden	12.83	Libya	0.46	Lithuania	4.78	Libya	0.029
Switzerland	12.78	Gambia, The	0.37	Estonia	4.66	Saudi Arabia	0.025
Cyprus	12.49	Saudi Arabia	0.36	Kazakhstan	4.60	Maldives	0.025
Greece	12.36	Algeria	0.35	El Salvador	4.36	Algeria	0.024
Norway	12.29	Indonesia	0.35	Ukraine	4.35	Guam	0.023
Netherlands	12.11	Tunisia	0.34	Belarus	4.35	Macao, China	0.022
Israel	12.10	Bangladesh	0.31	Reunion	4.12	Jordan	0.022
Hong Kong	12.04	Kiribati	0.31	Kyrgyz Republic	3.59	Iran, Islamic Rep.	0.020
...	...	...	...	...	...	...	...
Rwanda	-18.44	Kazakhstan	-0.31	Papua New Guinea	-3.23	Liberia	-0.028
Mozambique	-19.17	Russia	-0.31	Uganda	-3.40	Central African Rep.	-0.032
Somalia	-19.19	Burundi	-0.31	Comoros	-3.40	Malawi	-0.033
Niger	-19.48	Ukraine	-0.31	Pakistan	-3.43	Burundi	-0.036
Ethiopia	-19.55	Belarus	-0.34	Guinea	-3.64	Mozambique	-0.037
Angola	-19.66	Rwanda	-0.46	Afghanistan	-4.63	Zimbabwe	-0.038
Guinea	-19.77	Zambia	-0.46	India	-4.68	Botswana	-0.038
Afghanistan	-20.25	Uganda	-0.47	Bangladesh	-5.25	Zambia	-0.041
Guinea-Bissau	-21.43	Zimbabwe	-0.47	Maldives	-6.11	Uganda	-0.041
Sierra Leone	-25.74	Botswana	-0.55	Nepal	-6.21	Rwanda	-0.045

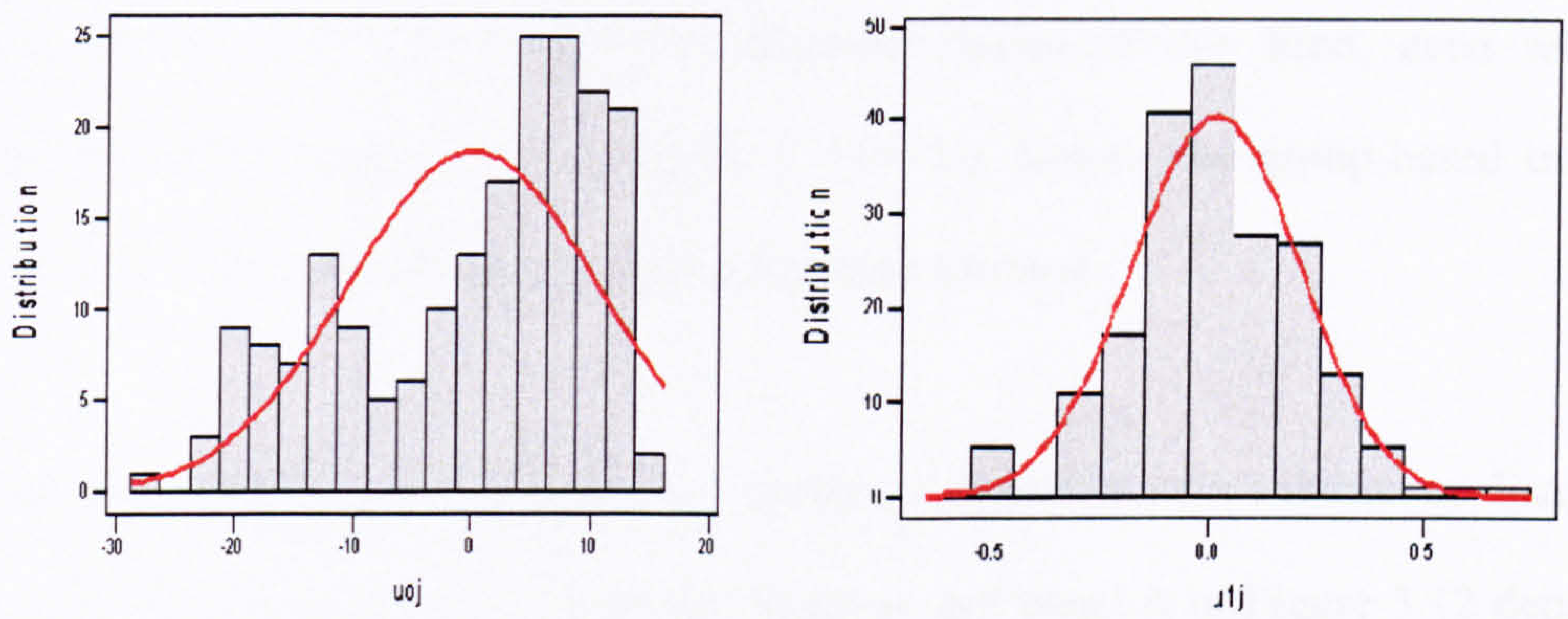
3.3.2 Group-based semi-parametric modelling

A disadvantage of the continuous random-effects approach is that it makes quite strong assumptions about the multivariate normality of the country effects. We can see from the histogram in Figure 3.11, which shows the country residuals for the combined male and female model, that while the u_{1j}, the differential slopes, are approximately normally distributed, the u_{oj}, the differential intercepts are negatively skewed in comparison to a normal curve with the same mean and standard deviation. There are relatively few countries that worse than the global average. An attractive alternative is therefore to work with a discrete unspecified distribution for the country effects which makes less strong distributional assumptions. In many settings as here, it is more natural to classify groups into a small number of types of countries than to estimate them as a continuous distribution.

The remainder of this section is organised into three parts. The first clarifies the underlying statistical method of the group-based trajectory model, including the theoretical background of the model structure, the equations and the hypothesis testing. The second part shows how this technique can be used to find underlying trajectories with a simulation dataset of know properties. The third section sets out the strategies that will be applied for answering the aims of this study.

For modelling developmental trajectories, hierarchical modelling (that is multilevel modelling; Bryk and Raudenbush 1987, 1992; Goldstein 1995) and latent growth curve modelling (Meredith and Tisak 1990; Muthen 1989; Willett and Sayer 1994) are two main branches of methodology. A third, alternative method – group-based trajectory modelling – was suggested by Nagin and Land in 1993.

Figure 3.11 Testing the validity of the Normality assumptions contained in the multilevel model: (left) for the random intercepts (μ_{0j}); and (right) for the random slopes (μ_{1j})

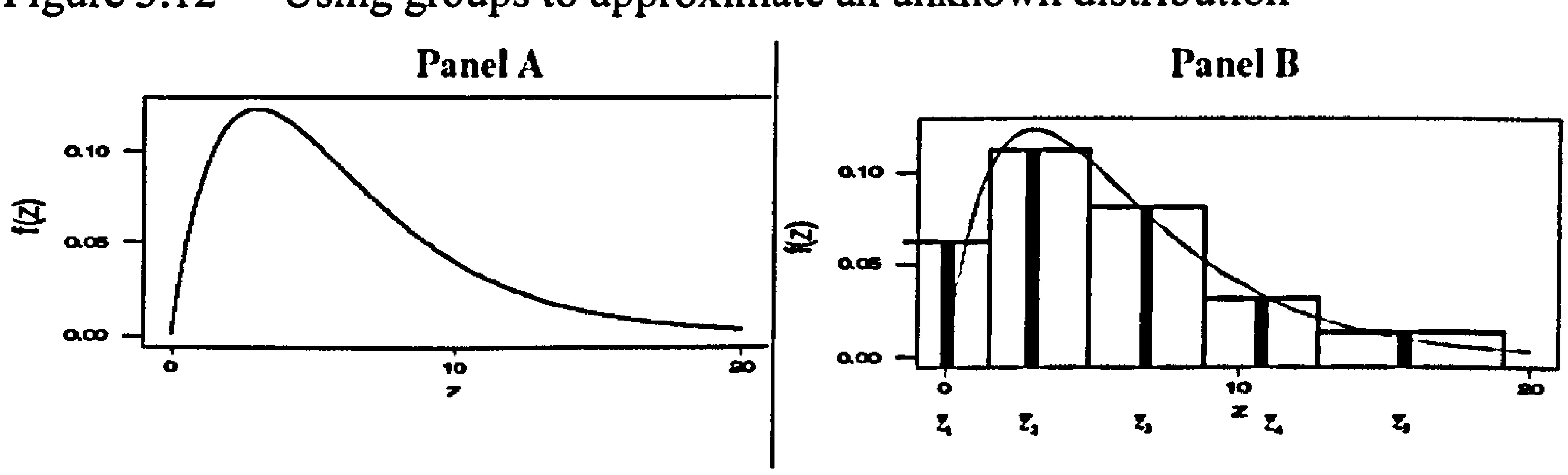


These three methods are all designed to measure and explain differences across population members in their developmental processes. They share the common goal of modelling heterogeneity in developmental trajectories but each makes different technical assumptions about the distribution of trajectories in the population. It is these assumptions that distinguish the three approaches. The two former approaches both model the population distribution of trajectories based on continuous distribution functions and therefore assume that the parameters are distributed throughout the population according to the multivariate normal distribution. However, the group-based trajectory modelling takes a qualitatively different approach to modelling individual differences. Rather than assuming that the population distribution of trajectories varies continuously across individuals and in a fashion that can ultimately be explained by a particular multivariate distribution of population parameters (usually normal), it assumes that there may be clusters or groupings of distinctive developmental trajectories that they may reflect distinctive aetiologies.

This mixture model is designed to take the unobserved heterogeneity within a population into account through the modelling. As noted by Pickles and Angold (2003, p541) "both theoretical and empirical work in statistics has shown that the nonparametric estimator of the underlying distribution, essentially the best fitting distribution, is just such a set of discrete classes of this kind, even when the underlying distribution is continuous ". For this reason, the group-based trajectory method is often described as a semi-parametric method.

The idea of using a finite number of groups to approximate a continuous distribution is easily illustrated with an example. Suppose that panel A in Figure 3.12 depicts the population distribution of some behaviour Z. In panel B, this same distribution is replicated and overlaid with a histogram that approximates its shape. Panel B illustrates that any continuous distribution with finite end points can be approximated by a discrete distribution (i.e., a histogram) or alternatively by a finite number of "points of support" or 'mass points' (i.e., the dark shaded "pillars"). A higher number of support points yields a discrete distribution which more closely approximates the true continuous distribution. However, simulation evidence reported in Brame et al., (forthcoming) and Nagin (2005) suggests that relatively few points of support are required to approximate reasonably even complex continuous distributions of trajectories.

Figure 3.12 Using groups to approximate an unknown distribution



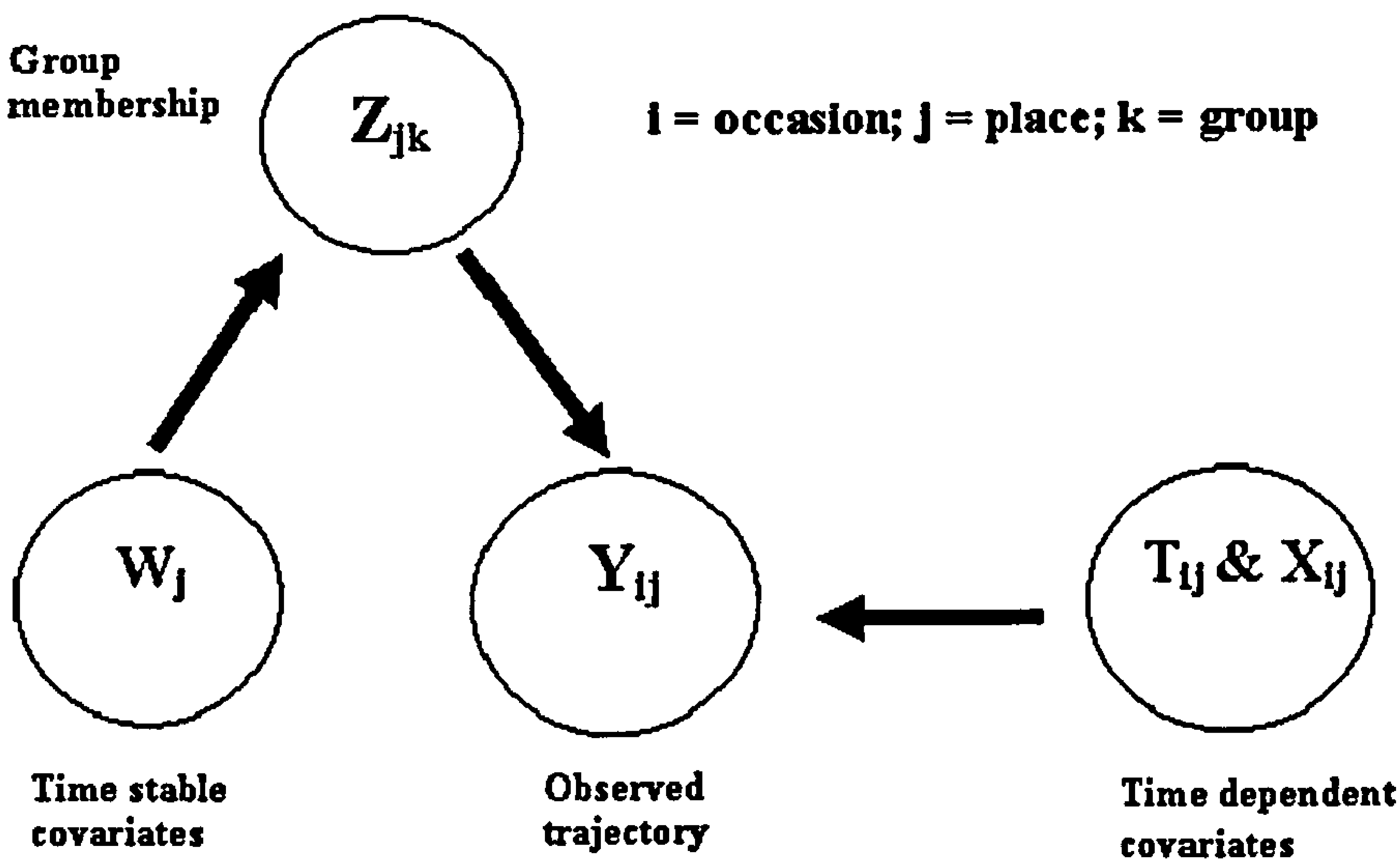
(Source: adapted from Nagin and Land, 1993)

3.3.2.1 Structure of the model

The general structure of the model is as illustrated in Figure 3.13 as a directed acyclic graph where arrows indicate the relationship of the outcome variable as a function of each variable in the model. Y_{ij} is the observed life expectancy for occasion i nested within place j. These depend on the time variable T_{ij} (usually as a set of polynomials) as well as time-dependent predictor variables X_{ij} which could be measures of inequality and GDP. Group membership, Z_{jk}, is defined by whether country j belongs to any group k. In practice this is the probability that a country belongs to a specific group. The final set of variables, W_j are the time-stable characteristics (or a risk factor) that potentially account for group membership.

The model is appropriate for data with average values changing smoothly as a function of the polynomial of time. Sharp changes can be handled through the inclusion of time dependent covariates with dummy variables representing regime shift.

Figure 3.13 Directed acyclic graph of the group trajectory model



More formally, a growth trajectory model with third order (i.e. cubic, t[5]) polynomial of time can be written as follows:

$$Y_{ij} = \sum_{k}^{G} (\beta_0^{(k)} Z_{jk} + \beta_1^{(k)} t_{ij} Z_{jk} + \beta_2^{(k)} t_{ij}^2 Z_{jk} + \beta_3^{(k)} t_{ij}^3 Z_{jk}) + \mu_j^{(k)} + \varepsilon_{ij} \qquad (3.1)$$

$$\mu_j^{(k)} \sim N(0, \sigma_\mu^{2(k)})$$

$$\varepsilon_{ij} \sim N(0, \sigma_e^2)$$

$$Z_{jk} = \begin{cases} 1 & if \quad g_j = k \\ 0 & if \quad g_j \neq k \end{cases}$$

$$g_j \sim multinomial(\pi_j, 1)$$

$$\log\left(\frac{\pi_j^{(k)}}{\pi_j^{(G)}}\right) = \beta_0^{(k)} + \beta_1^{(k)} w_j + \beta_2^{(k)} w_j + \beta_3^{(k)} w_j \qquad k = 1, \ldots, G-1 \qquad (3.2)$$

In equation 3.1 the observed life expectancy in country j at time i is related to a third order polynomial of time. The βs in this model are regression coefficients which give respectively the linear, quadratic and cubic relations between expectancy and time. The superscript k indexes the unknown groups and we can have one to G of these, each with a potentially different set of estimated β terms with distinctive trends. In this first equation there are two random terms which summarise the unexplained variation after the trends have been extracted. The μ_j are the between-country residual differences, while ε_{ij} are the within-country, between-occasion residual term. Assuming a normal distribution with mean 0, they can be summarised respectively in variance terms $\sigma_\mu^{2(k)}$ and σ_e^2. The formulation adopted here allows each group of places to have a different degree of residual variability and that is why the variance terms has the superscript k.

[5] It is common practice in the latent growth literature for a third order polynomial to be used as this is complex enough to capture major trends but not too complex that it compromises the ability to fit models reliably with convergence.

The latent group membership indicator, Z_{jk} is presumed to come from a multinomial distribution so in effect we are modelling the underlying probability (π_j) that a place belongs to a particular group. In the second equation this group membership is related to the time independent variables, W_{1j} etc, through a logit link in a multinomial model. Consequently these βs, when estimated, give the relationship between the log-odds of group membership and the characteristics of countries.

Procedures to fit this model have been implemented by Jones *et al.,* (2001) in their "PROC TRAJ" procedure which requires interfacing with SAS. Full details are available at http://www.andrew.cmu.edu/user/bjones/. This software supports three different distributions for the observed dependent variable (censored normal, Poisson, and Bernoulli). Here the censored normal is used to reflect the fact that life expectancy does not lie outside the zero to eighty-five range. This procedure should be appropriate for continuous data that are approximately conditionally normally distributed (Nagin and Tremblay 1999).

3.3.2.2 Model selection

There is still substantial debate over how to use test statistics to determine the number of groups to include in the model, as well as the specification of the order of the polynomial equation used to represent the shape of each group's trajectory (Ghosh and Sen 1985; Titterington et al., 1985). Here the recommendation of Nagin (2005) is adopted – to use a Bayesian Information Criterion, calculated as

$$BIC = \log(L) - 0.5k\log(N) + k$$

where L is the model's maximized likelihood, N is the sample size, and k is the total number of parameters in the model. The BIC is a 'badness' of fit measure which penalizes model complexity. Improved goodness of fit will be reflected in the likelihood; complexity is represented by the number of parameters which is determined by the order of the polynomial used to model each trajectory and the number of groups. D'Unger et al. (1998) have shown that the change in the BIC between models is an approximation to the log of the Bayes factor. The Bayes factor gives the posterior odds that the alternative hypothesis is correct when the prior probability that the alternative hypothesis is correct equals one-half. Muthen (2003) emphasises that the BIC is not really an index for addressing a model's absolute goodness of fit but, instead, is a relative fit measure comparing competing models.

In summary, model selection works inductively and in two stages. First the models are fitted without any time-independent variables but including a third order polynomial of time for a single group, and the BIC is calculated. Then two groups are specified and the BIC obtained. Next, three groups are fitted then four, five and so on until there is no further reduction in the BIC. This is then the most parsimonious model in terms of groups with distinctive trends. Once this has been identified, time-independent variables are included so as to account for group membership.

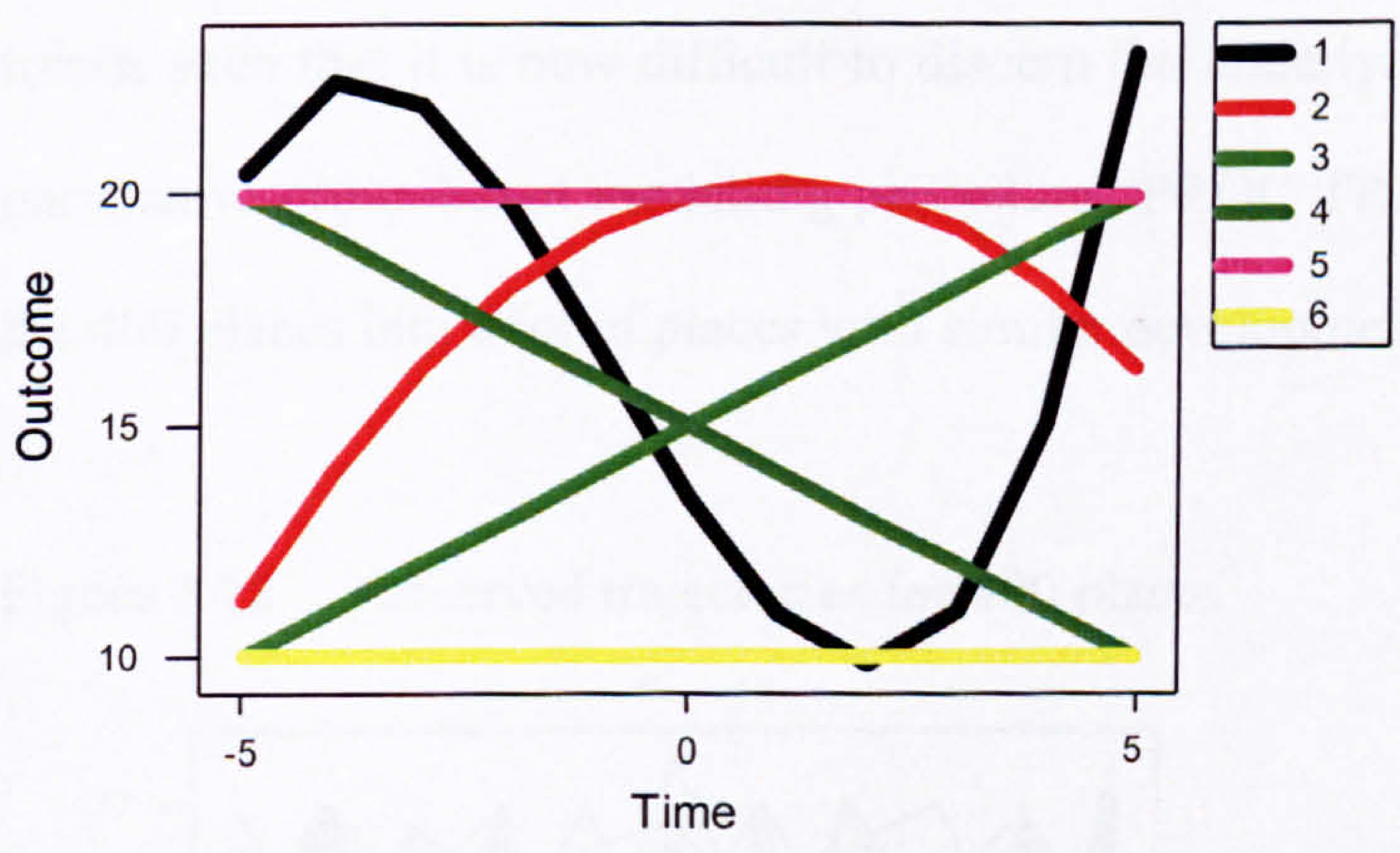
3.3.2.3 Data simulation

The following section uses simulated data in which the true model is known, to explore and to evaluate the capability of the group-based trajectory method to identify the distinctive features of a highly irregular but still continuous population distribution of trajectories. I have carried out a two stages of data simulation. Firstly, I created six different types of underlying trajectory in which the true model is known and added in level-one between occasion random noises. Then I added a level-two, between- places variation (of the random intercept form) with differing variability to each of the six distinctive trends. This simulation allows us to examine how the software implementation identifies different group trajectories as well as their level-one and -two variances.

First stage

The data set comprises 400 places of continuous outcome data which are under six types of developmental trajectory across eleven periods. They have been created in the software Minitab as follows. It is a data set with two levels structure of occasions repeatedly measured through time, and which are nested within places. The distinct six types of group trajectory are displayed in Figure 3.14 and the true generating equations are given Table 3.7.

Figure 3.14 Underlying polynomial trends with six types of place trajectories



Group one is a polynomial trend related to cubic time and represents 125 places. The black curve shows a strong trend with a clear 'up and down' cycle through time. There are 100 areas in Group 2 with an underlying non-linear one-bend trajectory related to quadratic time; this is shown in red. Two linear trends with same slope but opposite signs form the next two groups: the increasing pattern of Group 3 coloured in light green and the other, decreasing pattern of Group 4 coloured in dark green. Data have been generated for 50 places for each group. The final trajectories are both flat, with Group 5 being high (25 places, purple line) and Group 6 being low (50 places; yellow line)

Table 3.7 Detail equations information of six types of places

Type of Groups	No of Places (Group membership)	Trajectories
Group1: polynomial	125 (31.25%)	$13.5 - 3\times time + 0.33\times time^2 + 0.13\times time^3$
Group2: quadratic	100 (25.00%)	$20 + 0.5\times time - 0.25\times time^2$
Group3: linear	50 (12.50%)	$15 + 1\times time$
Group4: linear	50 (12.50%)	$15 - 1\times time$
Group5: high intercept	50 (12.50%)	20
Group6: low intercept	25 (6.25%)	10

Figure 3.15 shows the simulated data when level-1 random noise (with a variance of 3.6) is added to the underlying trends. This is quite a lot of noise to the signal of trends such that it is now difficult to discern the underlying patterns. The SAS semi-parametric group-based modelling procedure "PROC TRAJ" was now used to group the 400 places into a set of places with similar developmental trajectories.

Figure 3.15 Observed trajectories for 400 places

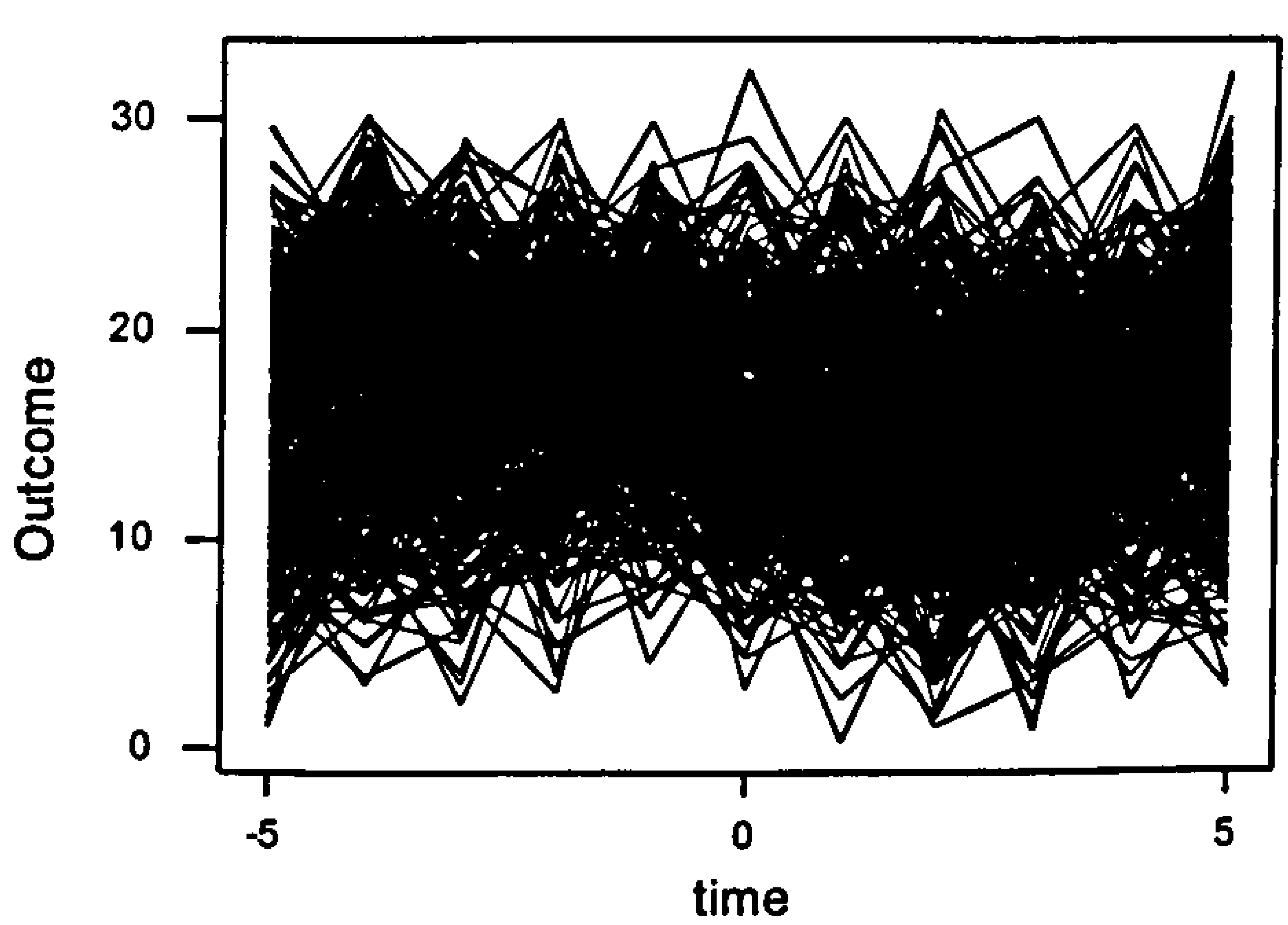


Table 3.8 gives the BIC results from fitting a sequence of models with one to seven groups. In fact two BIC statistics are given by the software; the first is on the basis that all the remaining variation lies between places so that there are only 400 independent observations to form the degrees of freedom. The second assumes that all the remaining variation is at level 1 between occasions so that there are some 4,400 independent observations (that is this BIC is based on the assumption of balanced data). Both type of BIC indicate correctly that there are six underlying groups with a 100% probability of a correct model being identified.

Having identified the 6 distinct trajectories, the models were refitted with only significant (at 0.05 level) polynomial trends included. Figure 3.16 plots the results with panel a showing the average of places in each group though time, while Figure 3.16b plots the predicted values based on the estimated polynomial trends.

Table 3.8 Both BIC result from SAS

Groups	BIC(400)	Diff[6]	Probability[7]	BIC(4400)	Diff	Probability
1	-13623.1		0	-13629.1		0
2	-13236.8	386.27	0	-13248.8	380.28	0
3	-12961.8	274.98	0	-12979.8	268.99	0
4	-12848.7	113.12	0	-12872.7	107.13	0
5	-12675.9	82.14	0	-12702.2	83.34	0
6	-12569.0	197.59	1	-12594.2	195.20	1
7	-12586.8	-17.8	0	-12619.2	-25	0

Table 3.9 gives the results in comparison to the values that were used to generate the original data. Clearly, the fitted trajectories closely approximate the true shapes and there is also a close correspondence for the percentage of places attributed to each group. Even the size of the level 1 random term is correctly estimated. The semi-parametric group-based trajectory method has demonstrated unequivocally its

[6] The change of BIC from the model itself to the previous model

[7] The probability of being the correct model

capability to capture the unobserved subgroups in the presence of considerable level-1 random variation.

Figure 3.16 Results of underlying polynomial trends with six types of place trajectories: (a) average of places in each group; and (b) predicted values based on the estimated polynomial trends

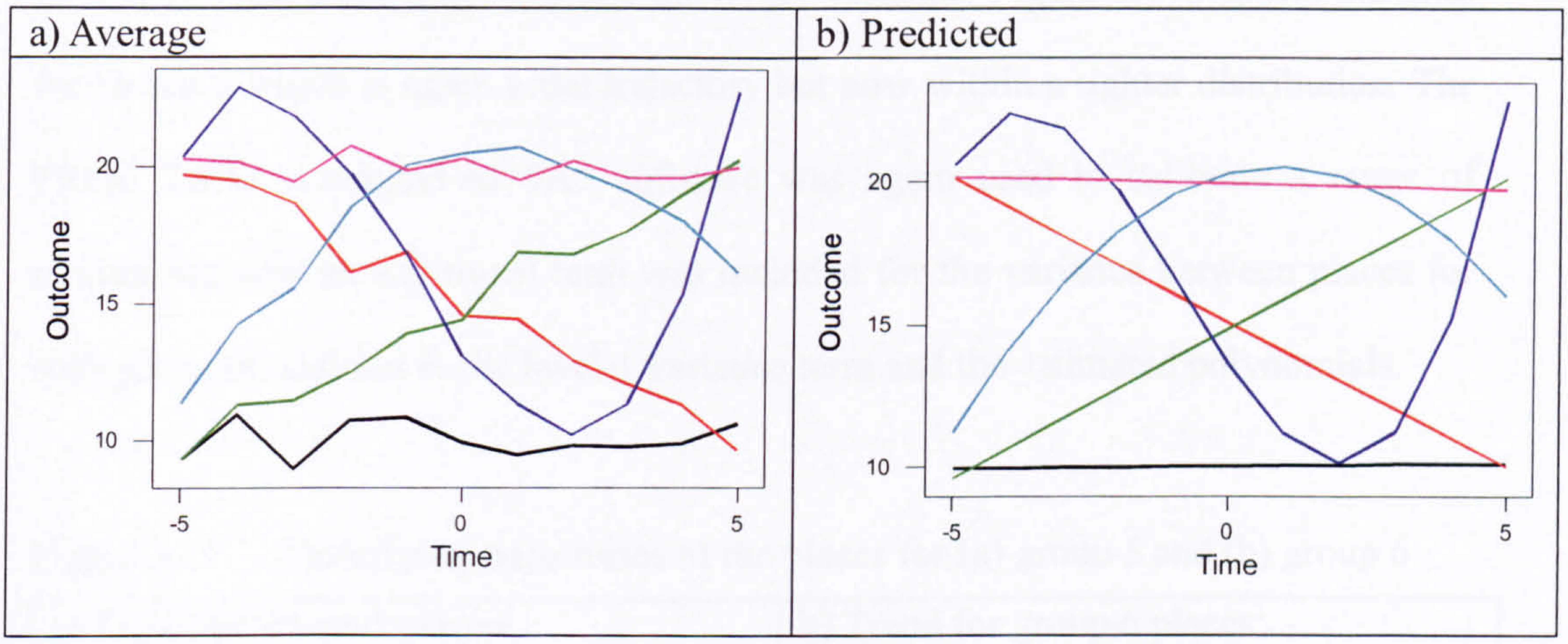


Table 3.9 Estimated results of 6-groups solution

Group	Parameter	Estimate (Truth)	Error	P value
1	Intercept	13.63 (13.5)	0.15	0.00
	Linear	-2.91 (-3)	0.08	0.00
	Quadratic	0.33 (0.33)	0.01	0.00
	Cubic	0.13 (0.13)	0.00	0.00
2	Intercept	20.34 (20)	0.19	0.00
	Linear	0.48 (0.5)	0.04	0.00
	Quadratic	-0.27 (-0.25)	0.01	0.00
3	Intercept	14.86 (15)	0.16	0.00
	Linear	1.05 (1)	0.05	0.00
4	Intercept	15.07 (15)	0.16	0.00
	Linear	-1.02 (-1)	0.05	0.00
5	Intercept	19.98 (20)	0.17	0.00
6	Intercept	10.01 (10)	0.23	0.00
	Level one variance	3.60 (3.6)	0.04	0.00
Group membership (%)		**Parameter (Truth)**	**Error**	**P-Value**
1		31.35 (31.25)	2.36	0.00
2		23.97 (25)	2.31	0.00
3		13.04 (12.5)	1.78	0.00
4		12.45 (12.5)	1.70	0.00
5		12.97 (12.5)	1.85	0.00
6		6.22 (6.25)	1.22	0.00

Second stage:

In the second stage of the simulation, I continued with the same dataset but now included extra random variation at the place level. In other words, I added level-two random variance of 1.2 to all of the groups, except for Group five which was given a larger variance of 1.8. The true relationships are shown in Figure 3.17a for Group five, which has a flat trajectory and the larger variance. Figure 3.17b shows the lines for Group 6 which is again a flat trajectory but now within a tighter distribution. The PROC TRAJ command in SAS software was again used to calibrate a range of models but now an additional term was included for the variance between places for each group in addition to the level-1 variance term and the estimated polynomials.

Figure 3.17 Underlying trajectories of the places for (a) group 5 and (b) group 6

a) Trend for group 5 places: Heterogeneous around flat, high trajectory	b) Trend for group 6 places: homogeneous around flat, low trajectory

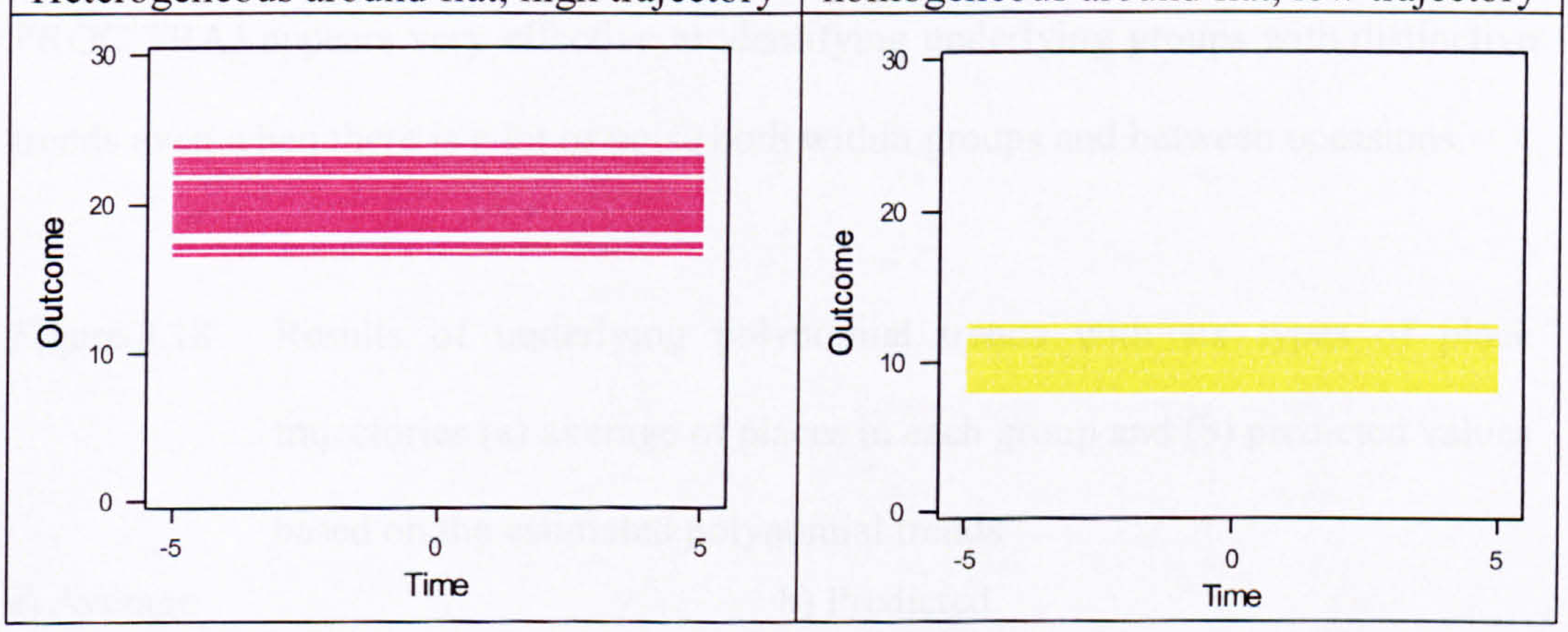


Table 3.10 presents the results for the BIC and again six groups are correctly identified. The graphs of Figure 3.18 show how well the models with only significant terms retained have captured the true trends of the simulated data. The relative size of group membership is well estimated (Table 3.11) but this is not as good as before (Table 3.9).

Table 3.10 Both BIC results for simulated data

Groups	BIC(400)	Diff	Probability correct model	BIC(4400)	Diff	Probability correct model
1	-13547.73	*	0	-13554.92	*	0
2	-13014.80	532.93	0	-13029.19	525.73	0
3	-12855.93	158.87	0	-12877.51	151.68	0
4	-12814.16	41.77	0	-12842.94	34.57	0
5	-12773.02	41.14	0	-12808.99	33.95	0
6	-12761.61	11.41	1	-12804.78	4.21	1
7	-12779.59	-17.98	0	-12829.94	-25.16	0

However there are some problems with the level-2 variances. While Group 5 is estimated correctly to have a large variance, an even larger one is found for Group 3. The variance for Group 6 is substantially underestimated. However, this is based on the smallest number of places. Moreover, the estimates clearly indicate by the size of their standard errors that there is a great deal of uncertainty about the estimates of the between group variance. The level 1 variance is very well estimated. In summary PROC TRAJ appears very effective at identifying underlying groups with distinctive trends even when there is a lot of noise both within groups and between occasions.

Figure 3.18 Results of underlying polynomial trends with six types of place trajectories (a) average of places in each group and (b) predicted values based on the estimated polynomial trends

a) Average b) Predicted

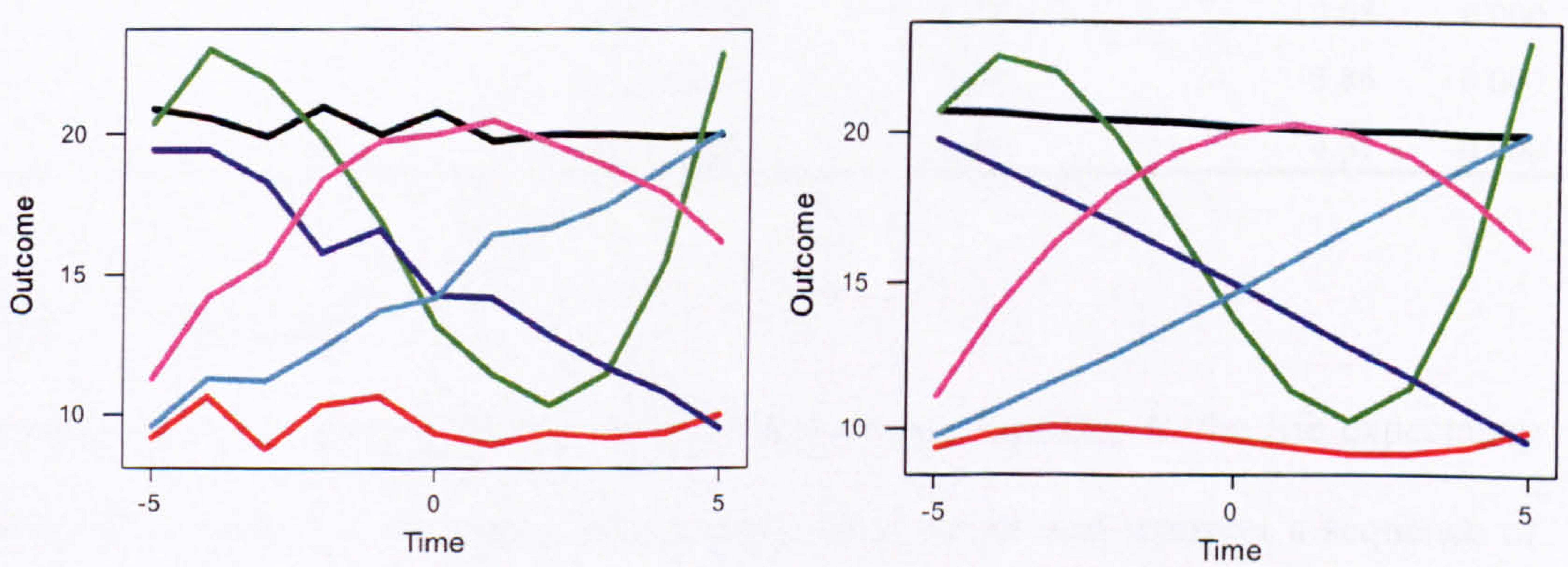


Table 3.11 Estimated results of 6-groups solution

Group	Parameter	Estimate (Actual)	Standard Error	T for H0: Parameter=0	Prob>|T|
1	Intercept	13.74 (13.5)	0.19	74.08	0.000
	Linear	-2.91 (-3.0)	0.08	-35.41	0.000
	Quadratic	0.33 (0.33)	0.01	28.40	0.000
	Cubic	0.13 (0.13)	0.00	29.66	0.000
2	Intercept	20.06 (20)	0.25	79.09	0.000
	Linear	0.44 (0.5)	0.10	4.67	0.000
	Quadratic	-0.26 (-0.25)	0.02	-17.20	0.000
3	Intercept	14.61 (15)	0.40	36.15	0.000
	Linear	1.06 (1)	0.14	7.57	0.000
4	Intercept	14.89 (15)	0.29	51.63	0.000
	Linear	-1.04 (-1)	0.14	-7.73	0.000
5	Intercept	20.22 (20)	0.41	49.35	0.000
6	Intercept	9.56 (10)	0.41	23.49	0.000
	Level one variance	3.61 (3.6)	0.04	86.80	0.000
1	Random intercept	1.17 (1.2)	0.14	8.35	0.000
2	Random intercept	1.36 (1.2)	0.26	5.32	0.000
3	Random intercept	1.62 (1.2)	0.28	5.82	0.000
4	Random intercept	0.96 (1.2)	0.27	3.57	0.000
5	Random intercept	1.50 (1.8)	0.28	5.32	0.000
6	Random intercept	0.52 (1.2)	0.50	1.04	0.298
Group Membership (%)					
1		31.29 (31.25)	2.38	13.15	0.000
2		25.96 (25)	2.77	9.39	0.000
3		12.84 (12.5)	2.02	6.37	0.000
4		12.19 (12.5)	1.72	7.08	0.000
5		12.25 (12.5)	2.09	5.86	0.000
6		5.47 (6.25)	1.19	4.61	0.000

3.4 RESULTS

In this section we apply the group-based trajectory procedure to the life expectancy data. The results are presented in two parts. First we fit and interpret a sequence of models that identified the underlying groups and their distinct trajectories. Second, having identified the groups we include inequality and GDP to predict group membership.

3.4.1 Modelling developmental trajectories

A sequence of models with one to eleven groups was fitted to the male plus female life expectancy data. The results for both type of BIC values (one is based on sample size in country level, the other is based on the number of occasions) and the probability of obtaining a 'correct' model are given in Table 3.12. Here I present each model in turn, discuss the distinctive trends and plot the groups geographically in Figure 3.19. This process gives a very good idea of world trends in life expectancy over thirty years. The maps are based on the group to which each country most probably belongs. The estimated percentage of countries in each is given in Table 3.13.

Table 3.12 Result for the changes in BIC values for one to eleven groups (both types)

Groups	BIC(196)	ΔBIC	Probability correct model	BIC(3260)	ΔBIC	Probability correct model
1	-8376.09	*	0	-8384.52	*	0
2	-7719.32	656.77	0	-7736.18	648.34	0
3	-7218.74	500.58	0	-7244.04	492.14	0
4	-6897.85	320.88	0	-6931.59	312.45	0
5	-6852.55	45.31	0	-6894.72	36.87	0
6	-6835.14	17.41	0	-6885.74	8.98	0
7	-6724.77	110.37	0	-6783.81	101.93	0
8	-6646.79	77.98	0	-6714.27	69.54	0
9	-6611.79	35.01	0	-6687.69	26.57	0
10	-6576.87	34.92	1	-6661.21	26.48	1
11	-6575.74	1.13		-6668.51	-7.30	

Model 1 (in Figure 3.19) shows the general global trend for the last three decades based on treating all countries as belonging to a single group. On average, life expectancy overall is 63.8 years in 1986 and has improved over time with a linear increase of 0.24 per year. In other words, every 4 years over the period, life expectancy has been raised by an average of 1 year. These results are identical to that produced by the fixed part of the multilevel model (Figure 3.7). The two-group solution on both BIC measures is a considerably better fit compared with the one group trajectory model. In one group are countries of the former Soviet Unions and Sub-Saharan Africa countries. This group is estimated to comprise about 21.4 percent of the sample population. Their distinctive trend is of a slightly lower life expectancy in the 1970s and a worsening experience in the 1990s. The other group consists of the rest of the world, including the wealthy economies, and represents over three-quarters of all countries. This group enjoys a slightly high life expectancy at the outset of the study period and continued improvement. Consequently, over time there has developed a substantial life expectancy gap between the two groups of some thirteen years.

Moving from two to three latent groups again brings an improvement in both BIC statistics. With this model, Group two includes North Africa, part of the Middle East, the Indian sub-continent, some South America countries and most of south Asia. These places experience a considerably better improvement over time with increasing life expectancy. In contrast, Group one has the worst experience of all, and this is mainly found in sub-Sahara Africa and in Eastern European countries. Groups 1 and 2 started the 1970s with similar life expectancies, but 30 years later there was a very substantial difference. Group 3 has 62 percent of all countries and includes Europe, Russia, the USA, Canada, Australia, China as well as much of Latin America and equatorial Africa. These countries have the highest Life expectancies but improvement over recent years has been relatively small.

It is worth stressing at this point, that the classification is non-hierarchical so that the new group has some members previously classified to Group one and some that were previously in Group two when only two latent classes were allowed. The geography of the classification remains however very marked. Of the four groups, the major change is that Russian and some central African countries from the former Group three, comprising of about 17.6 percent of all countries, now form a new group; This Group two, started with a relatively high life expectancy in the early 1970s with a little worsening over time.

For the five-group trajectory model, there is very little change; there is still a reduction in the BIC but this is not as great as before. In terms of the trajectories the only difference from previous model is that Libya and Oman, which both belonged to Group three in the last model, now form a new group, Group 1, with massive progress from a very low rate. With six latent groups there is an even smaller improvement in the fit, with the major change being that the former Group three has now split into new Groups three and five which have a 5 year gap in average life expectancy in 1970 but converge in the last decade of the twentieth century. Geographically, this model splits part of Africa, the Middle East and the Indian sub-continent into two groups.

The procedure of increasing the number of groups was continued until there were eleven latent classes. However, it was found that this number of groups was not a significant improvement over the ten group solution. Comparing the ten and eleven group solutions for both types of BIC (N=196) and BIC (N=3260), the approximation of the Bayes factor is 3.1, which represents moderate evidence for model eleven on Jeffrey's scale (Wasserman, 2000), and 6.7×10^{-4} which is strong evidence in favour of model ten. Thus, based on Jeffrey's scale as well as the BIC criteria, a ten-group model is the solution with the greatest support; it also meets our requirement of at least 1% of the countries is in each trajectory.

Table 3.13 Group prevalence for each trajectory

	Trajectory Group Prevalence (%)									
Model	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th
1	100.0									
2	22.4	77.6								
3	8.6	29.2	62.2							
4	4.1	17.6	27.4	50.9						
5	4.1	17.5	26.0	1.7	50.6					
6	4.1	16.7	23.7	0.0	13.7	41.8				
7	4.1	4.2	12.9	33.9	20.3	17.2	7.5			
8	4.1	4.2	12.1	24.0	20.4	7.1	21.9	6.2		
9	3.1	4.1	2.6	13.6	6.5	18.3	17.9	27.4	6.7	
10	3.1	2.6	2.6	14.0	2.1	24.4	14.1	27.8	6.4	2.9

For the preferred ten-group solution, Table 3.14 has the detailed-estimates and results on group membership for each and every trajectory. Group 8 had one of the highest life expectancies in 1970 and has shown continued improvement since them. Not surprisingly this includes Western and Northern Europe, the USA and Australia but also includes China and some South American countries. This trajectory has shown a 1 year increase in longevity every 5 years over the period. The second best life expectancy trajectory is experienced by Group 9. This group had a similar life expectancy with Group 8 but has seen no improvement over the period. Geographically, they are a very distinctive group in eastern European.

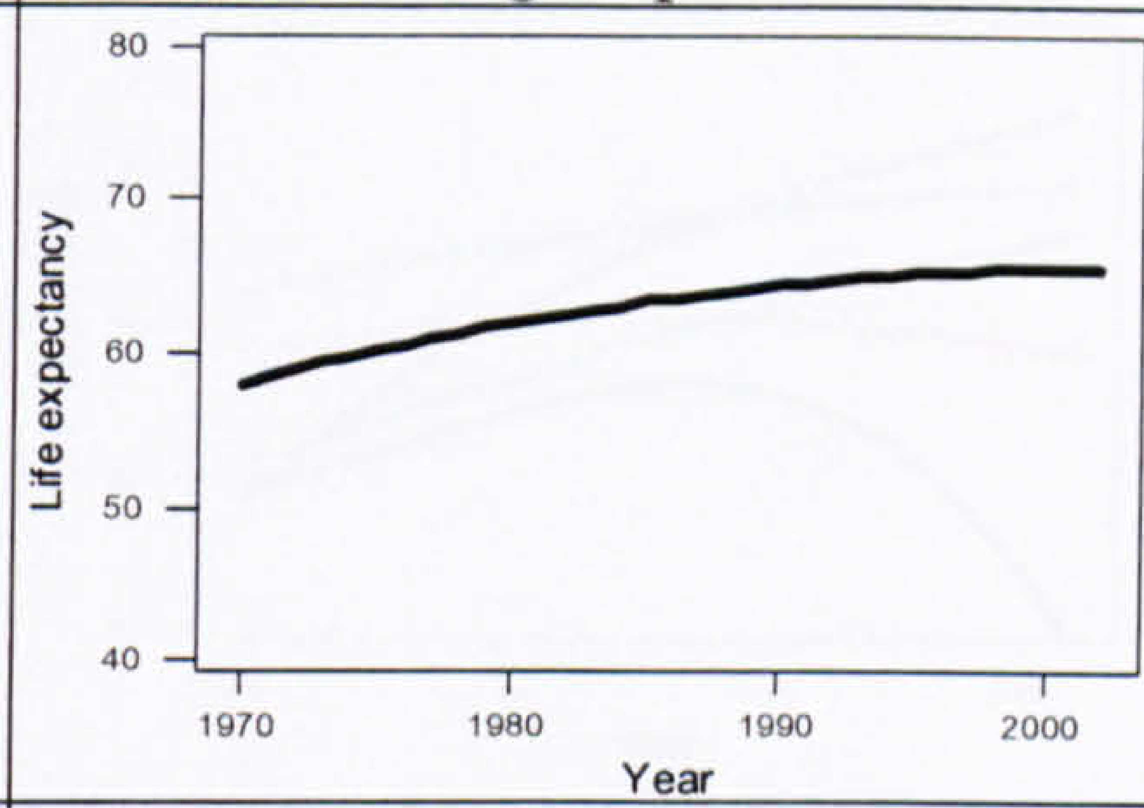
The worst life expectancy throughout the period is for Group 5. This consists of only two countries: Somalia and Afghanistan. They have the lowest life expectancy – around 40 years with some improvement over the thirty years. Another small but distinctive trajectory which has seen improvement, then reversal, and then improvement again is Group 3, which is made up of Tanzania, North Korea and Iraq.

Figure 3.19 Trajectories for life expectancy and their distribution geographically around the world under various grouping schemes.

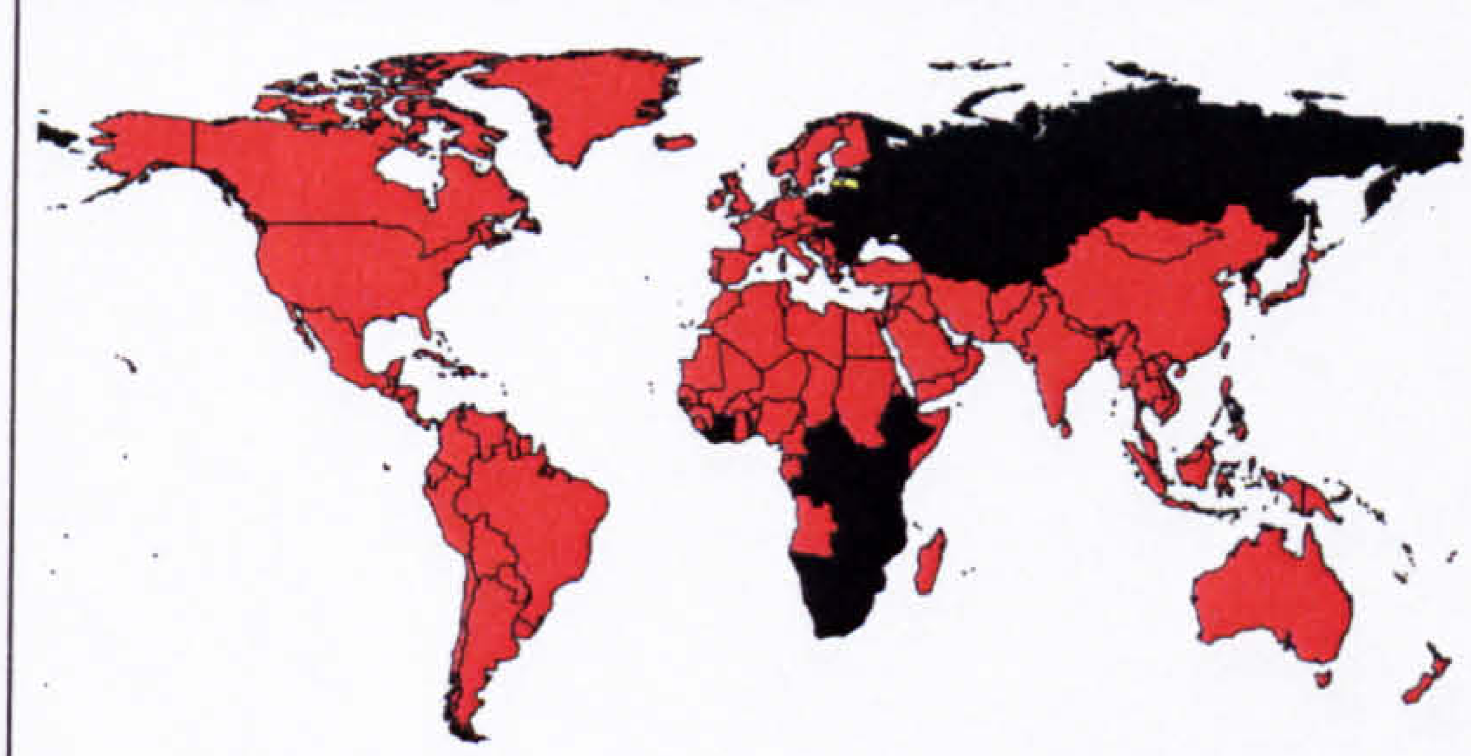
Map | **Prediction**

Model 1: one-group solution

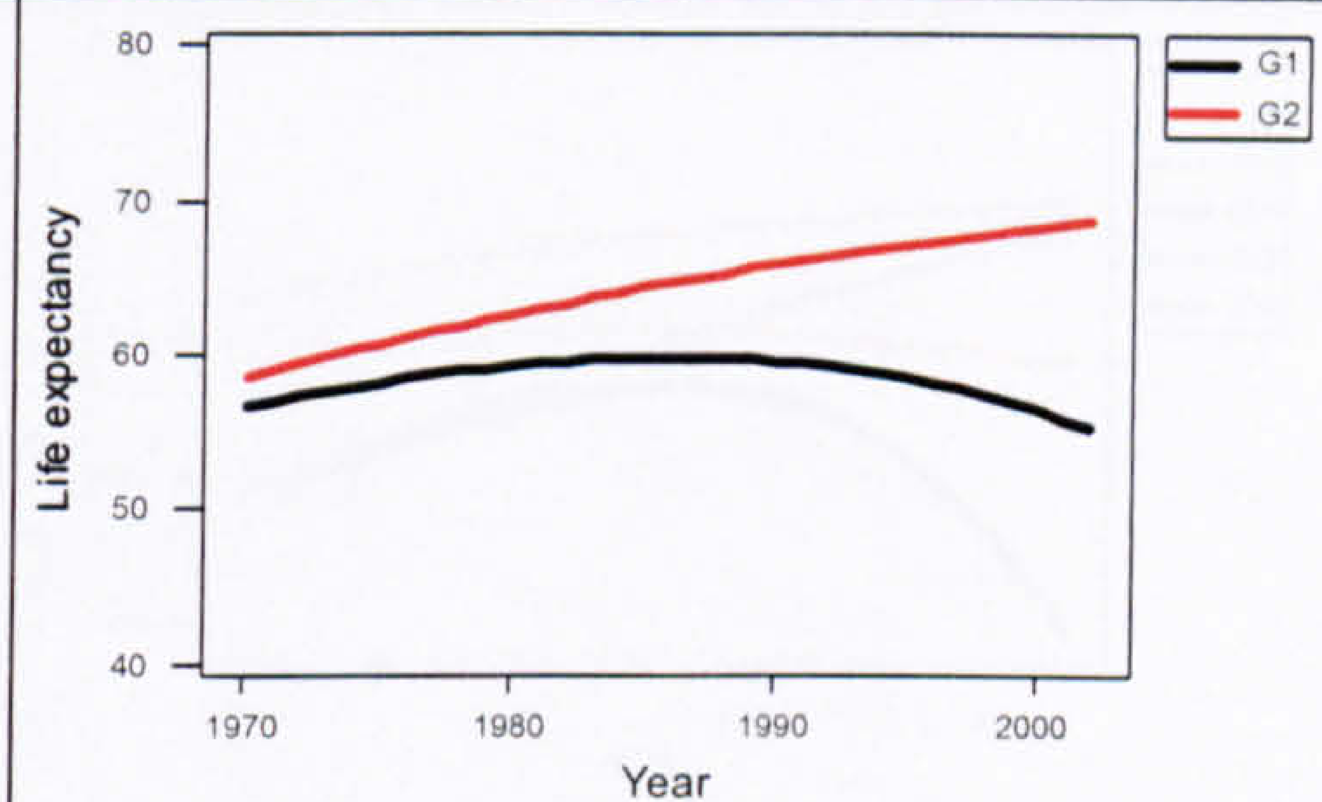
Model 1: one-group solution



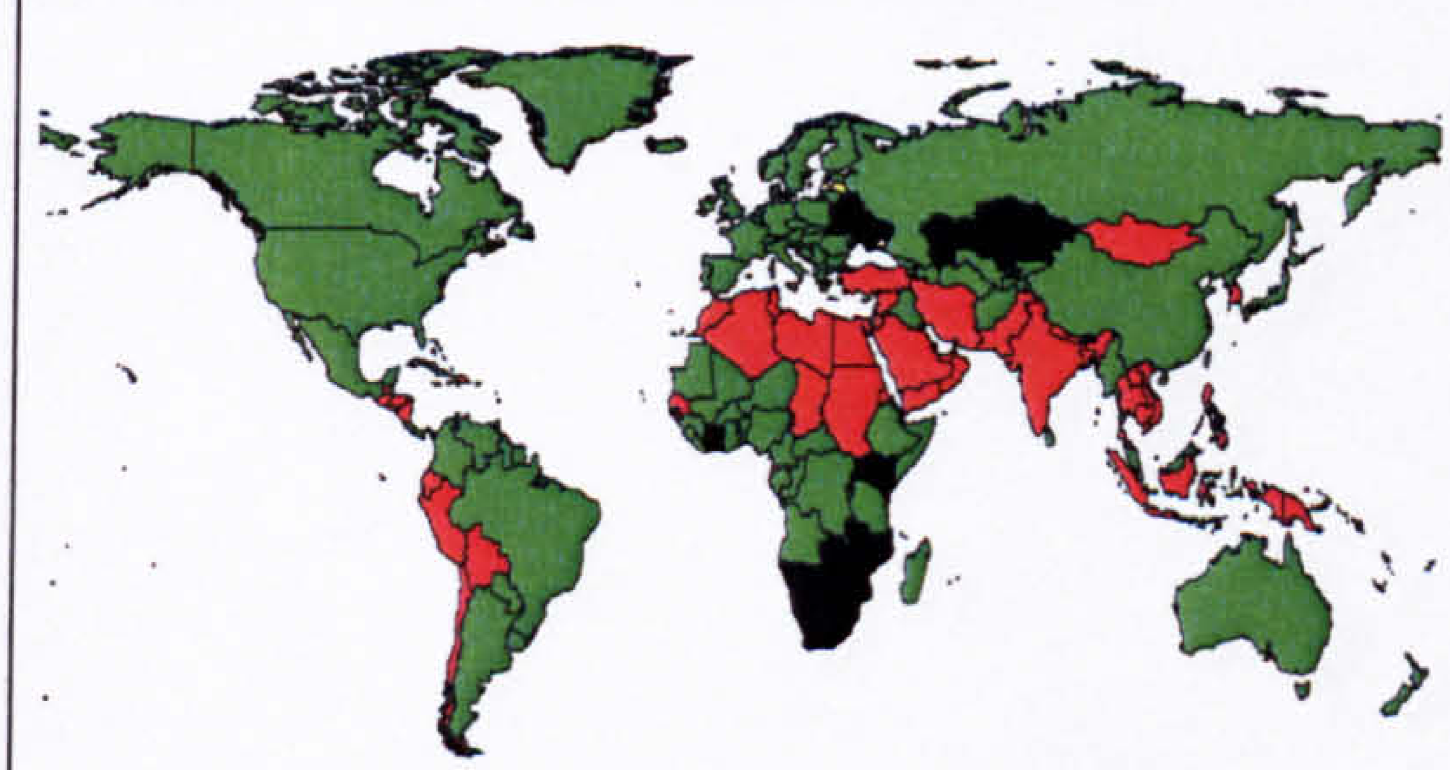
Model 2: two-group solution



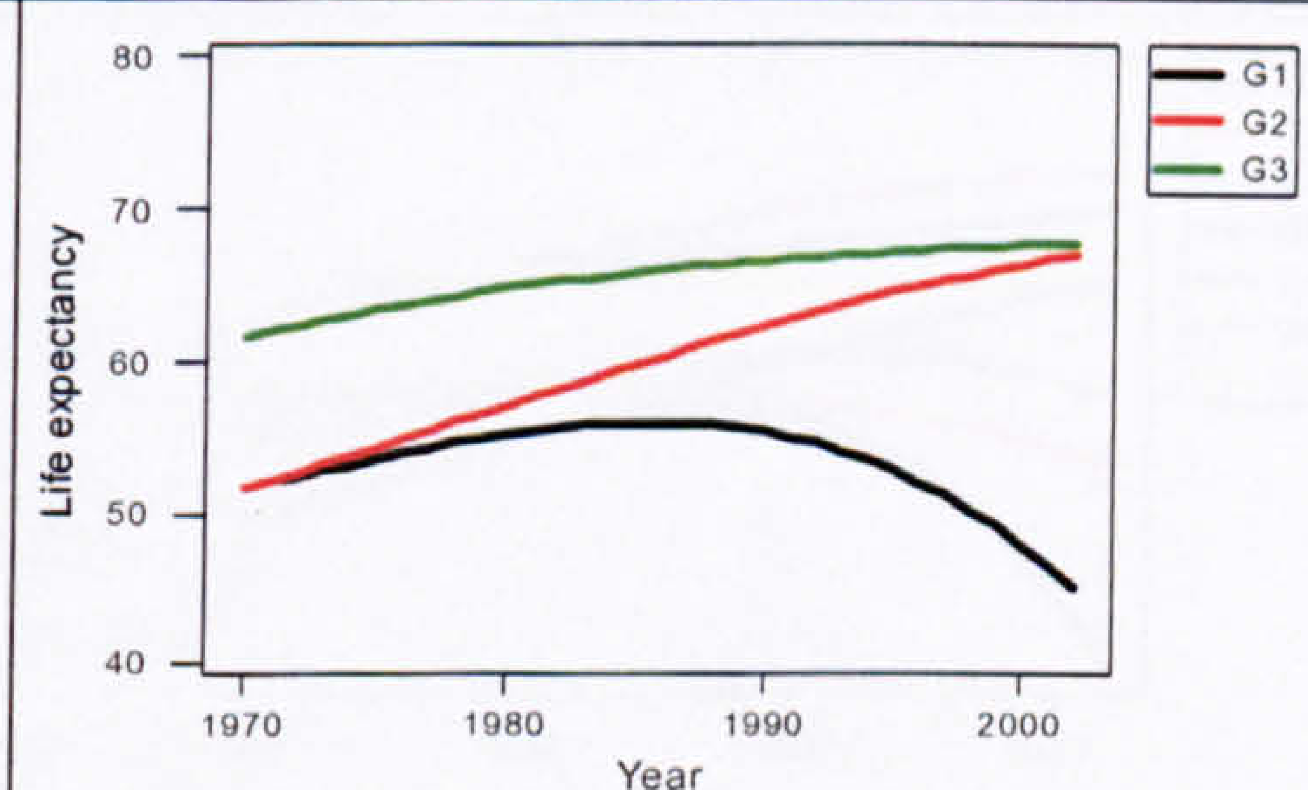
Model 2: two-group solution



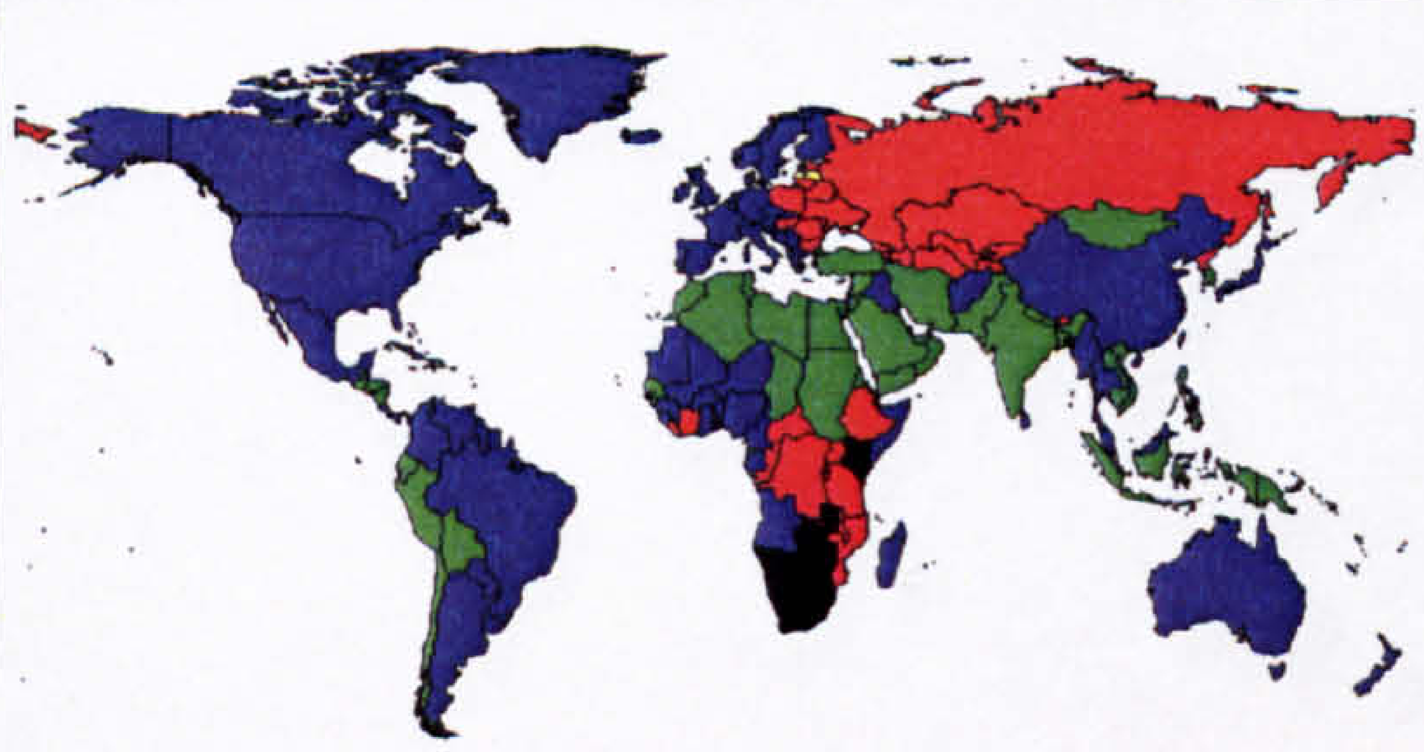
Model 3: three-group solution



Model 3: three-group solution



Model 4: four-group solution



Model 4: four-group solution

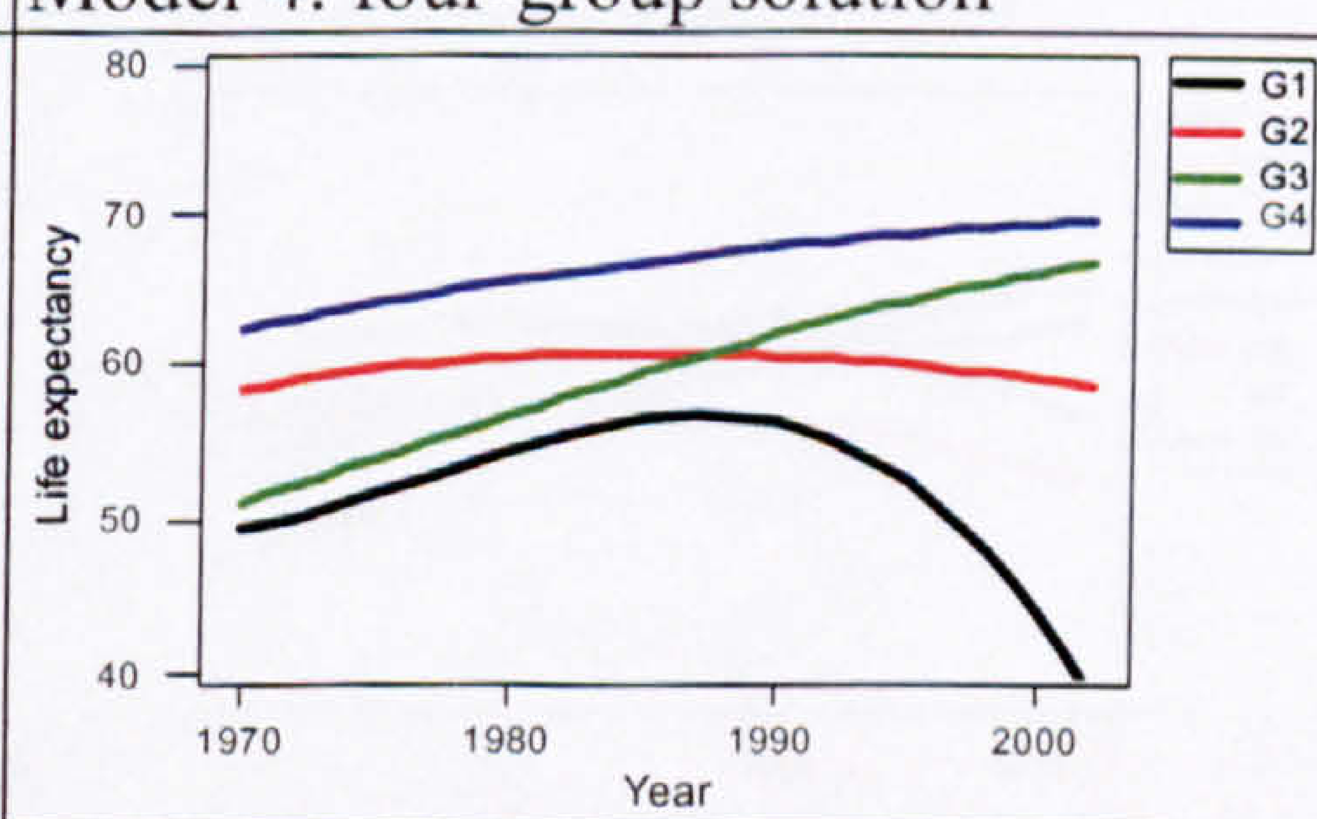


Figure 3.19 Trajectories for life expectancy and their distribution geographically around the world under various grouping schemes (continued)

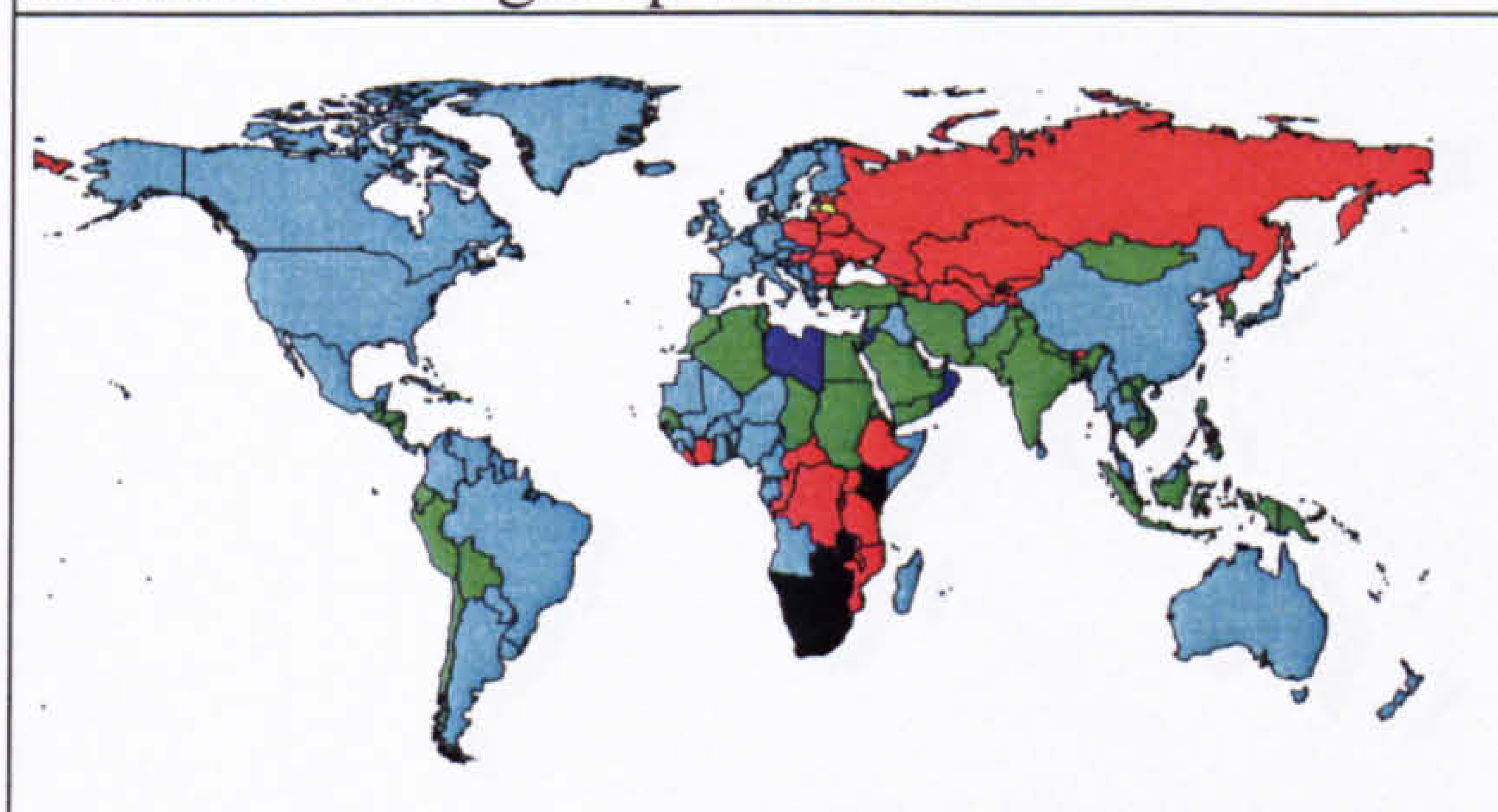
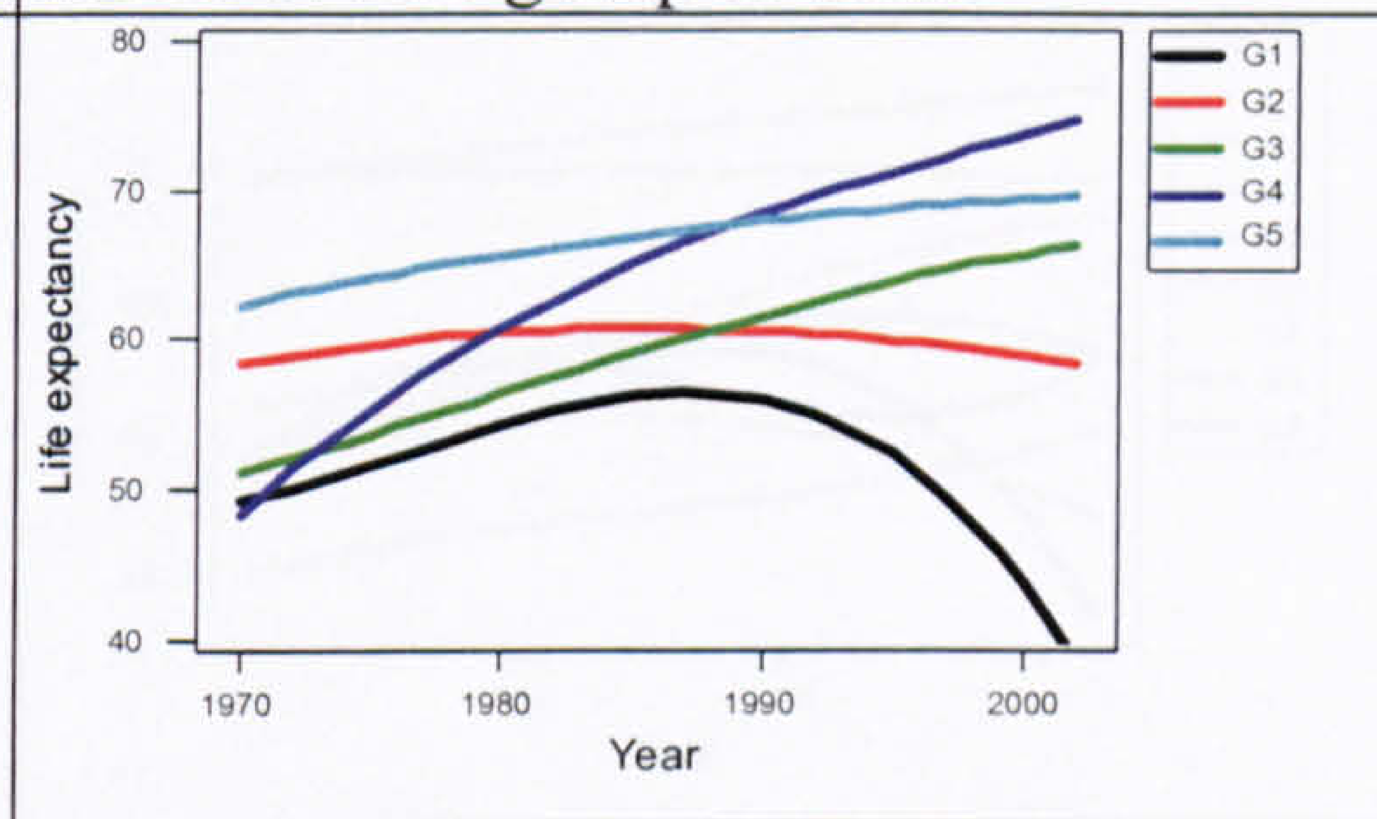
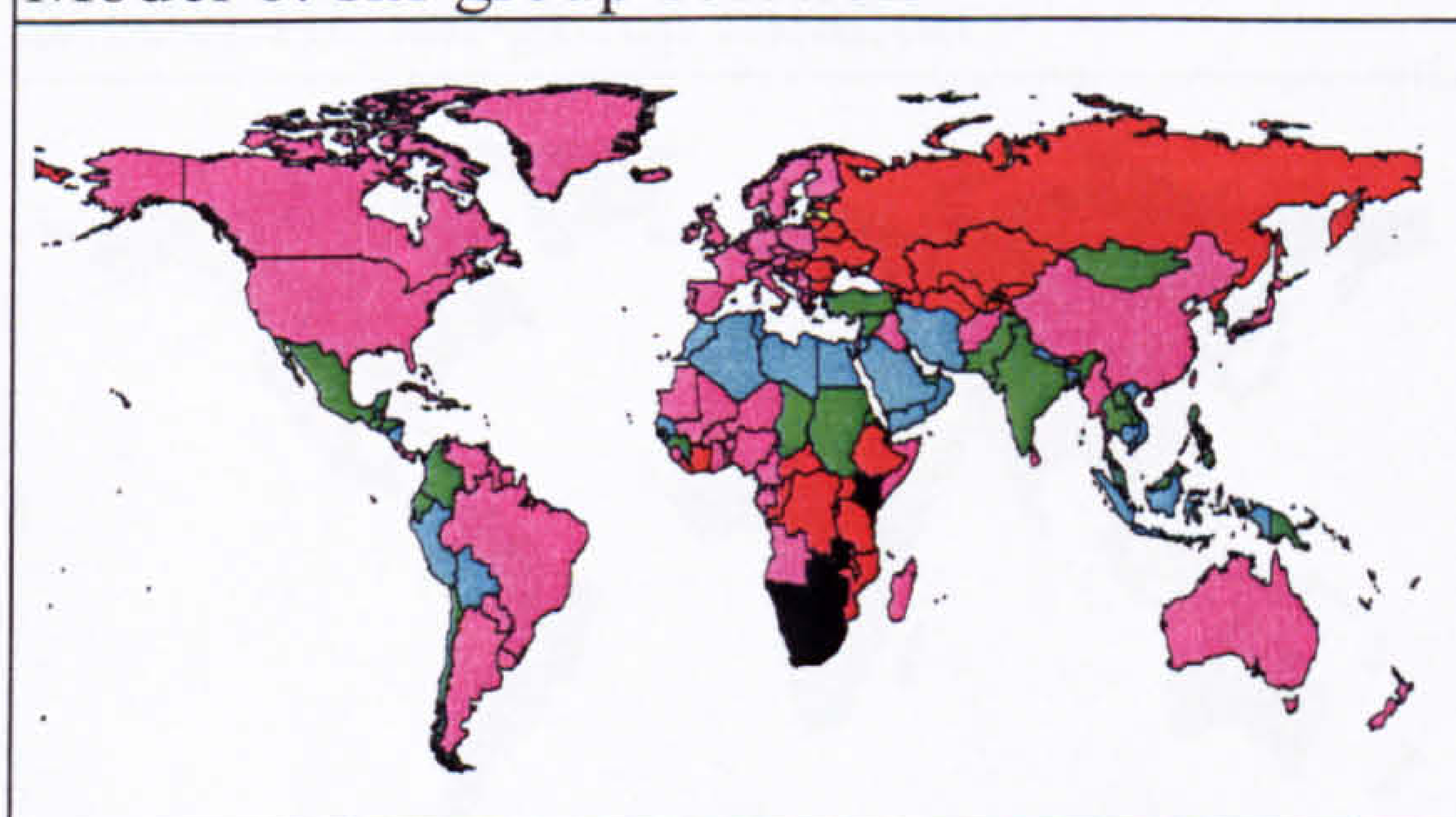
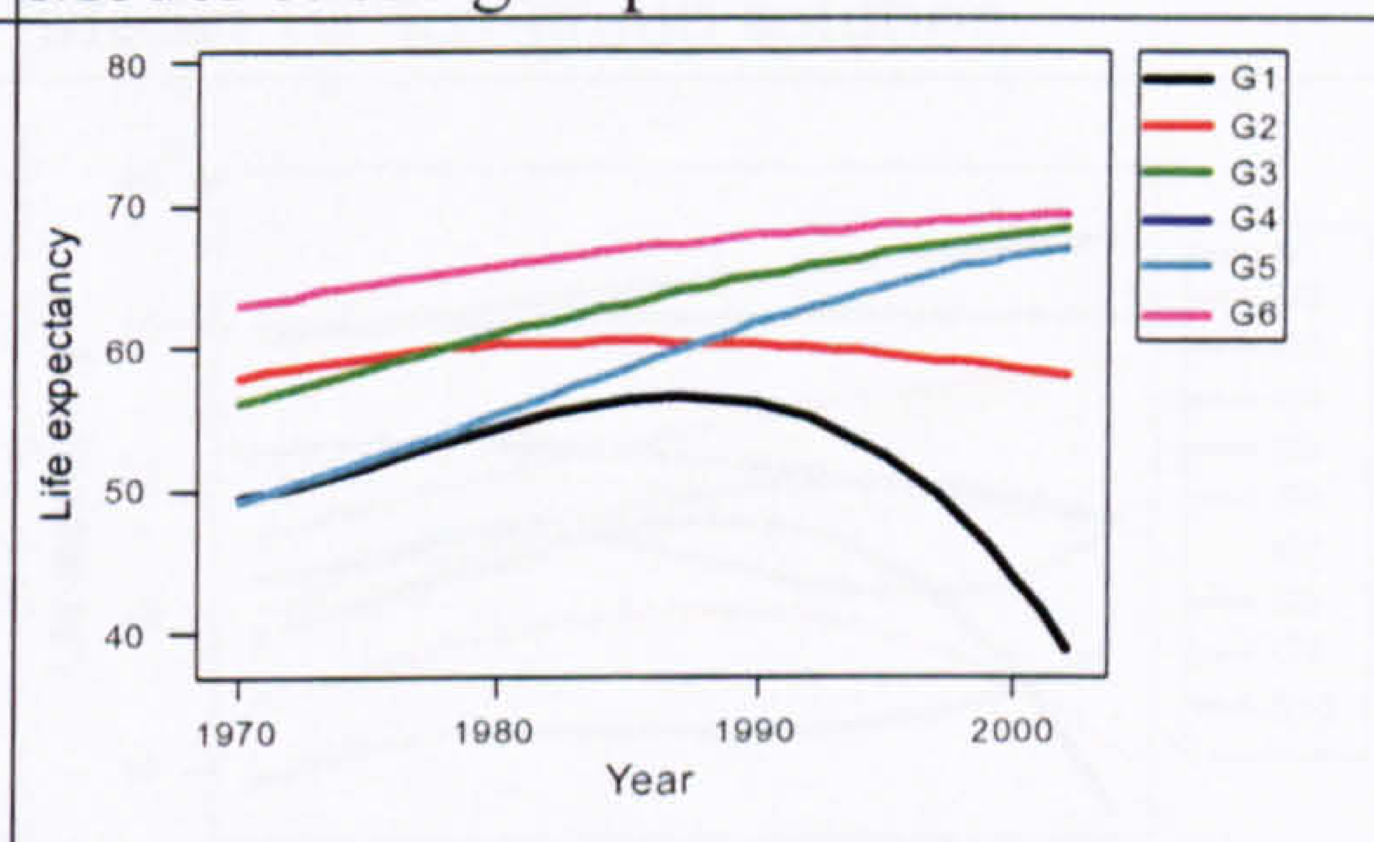
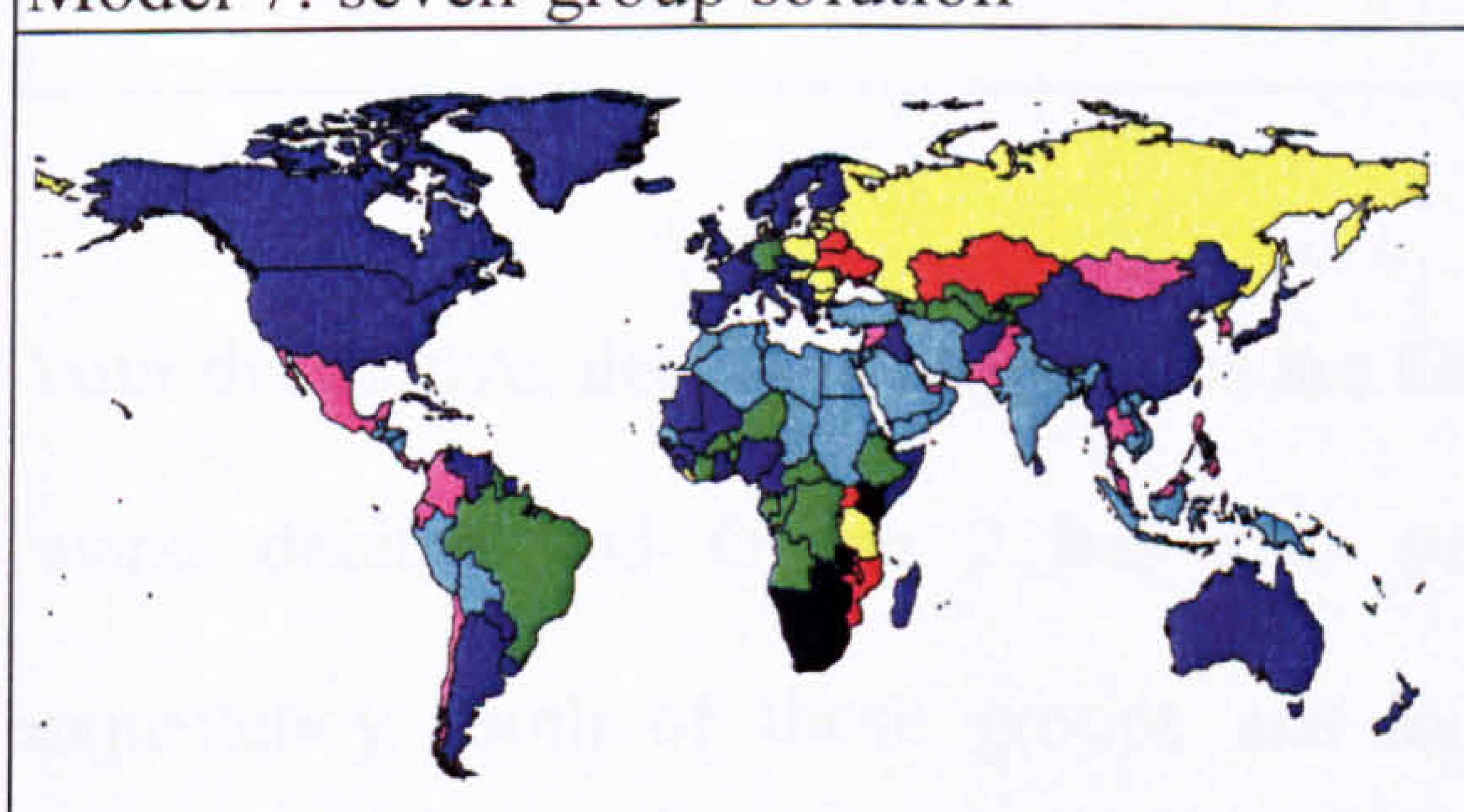
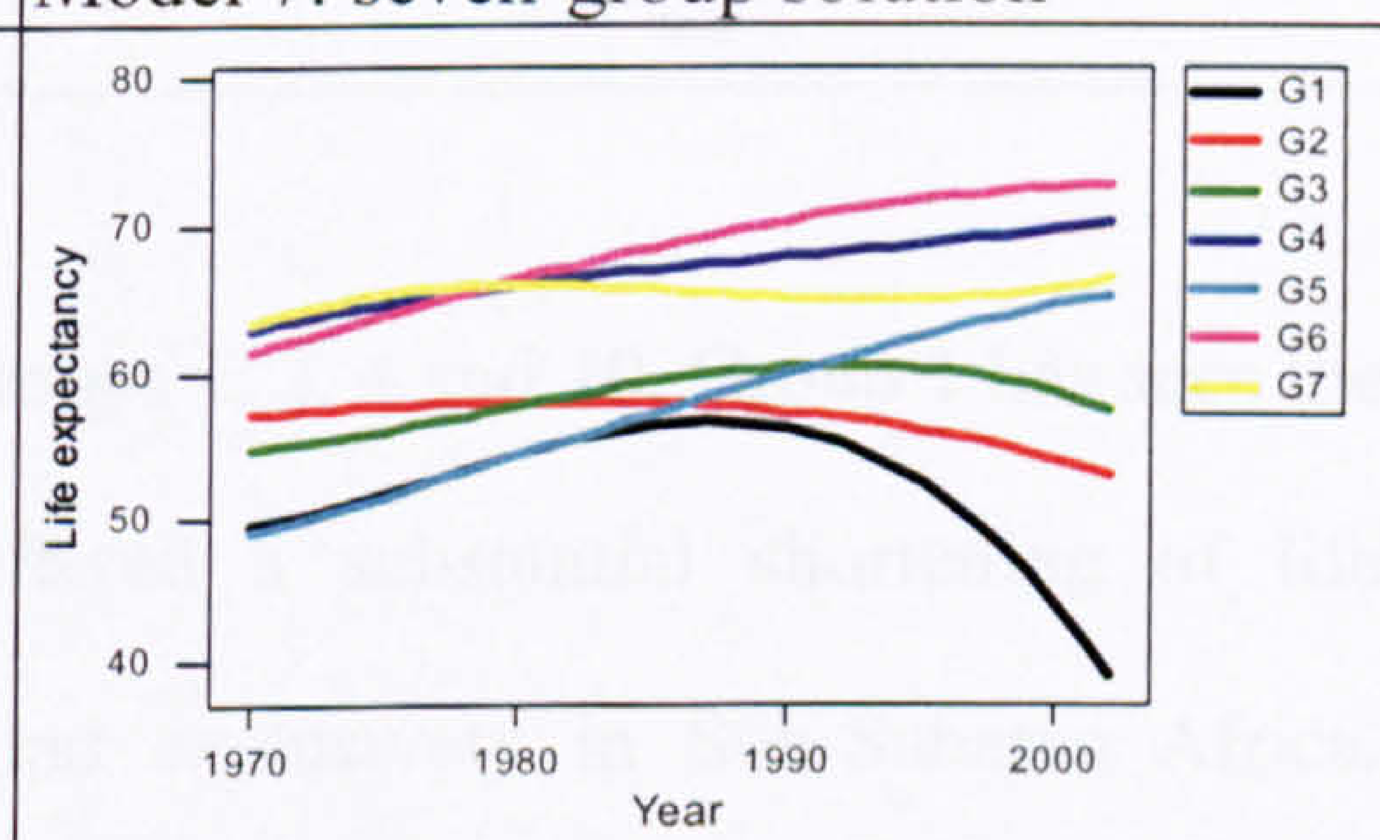
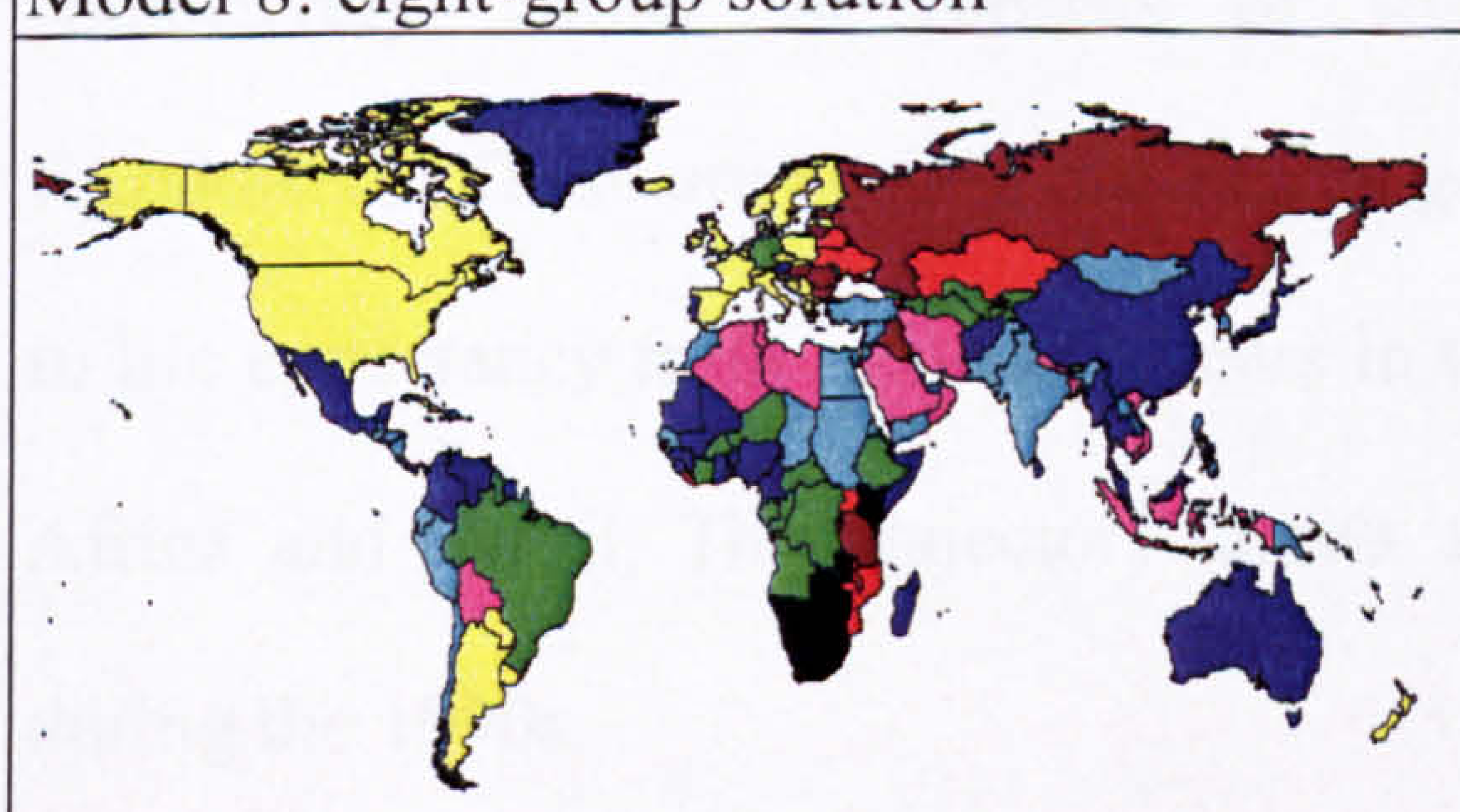
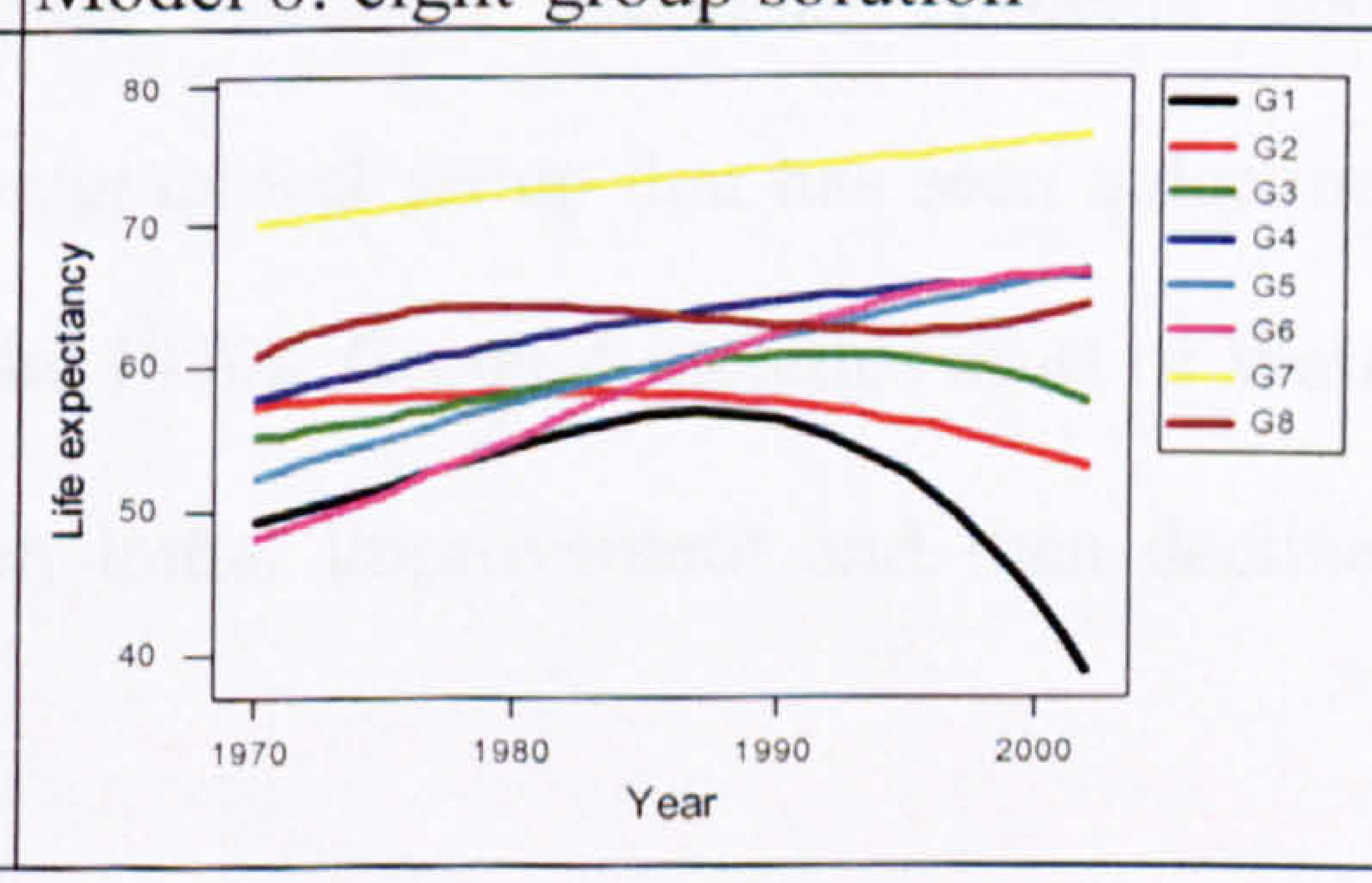
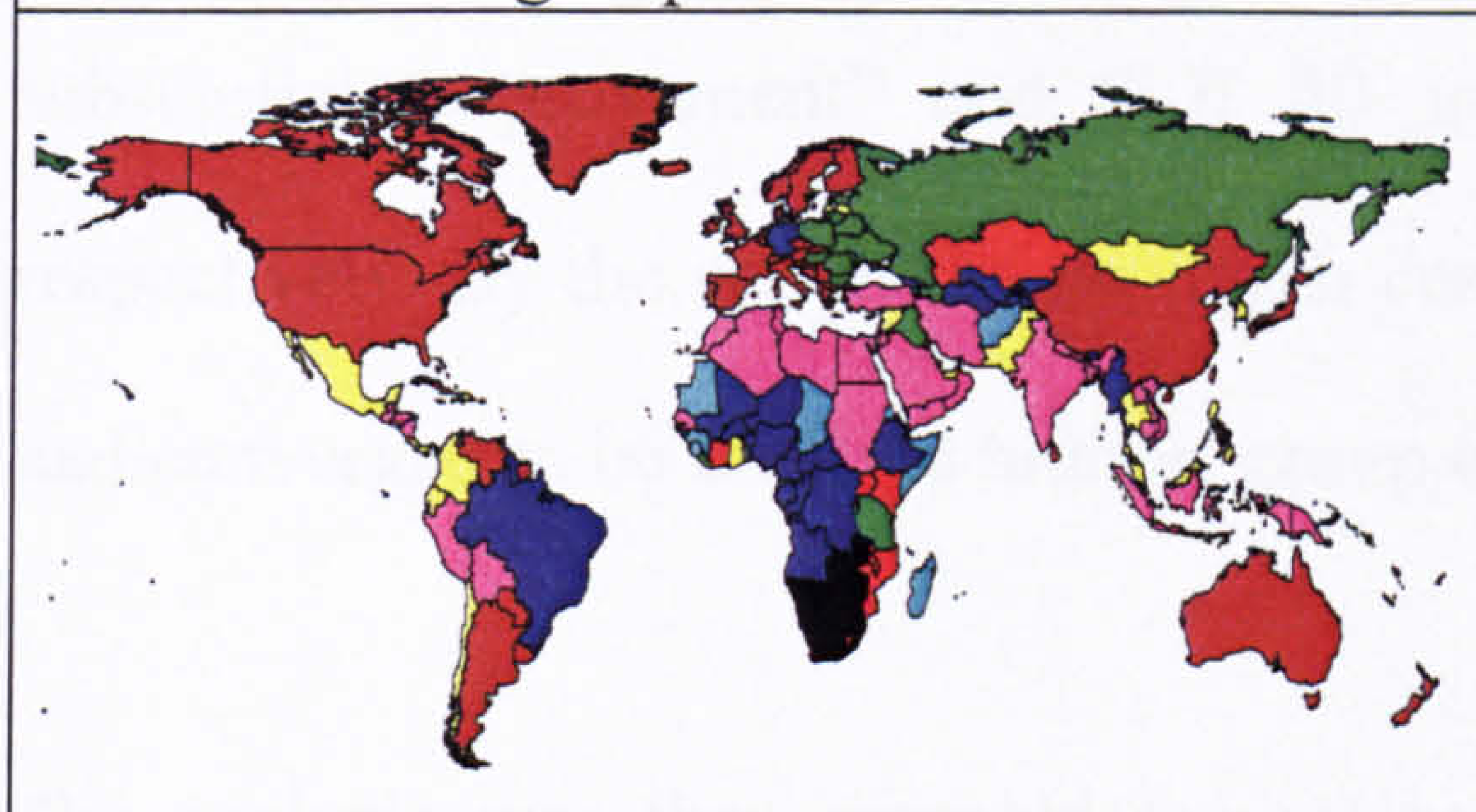
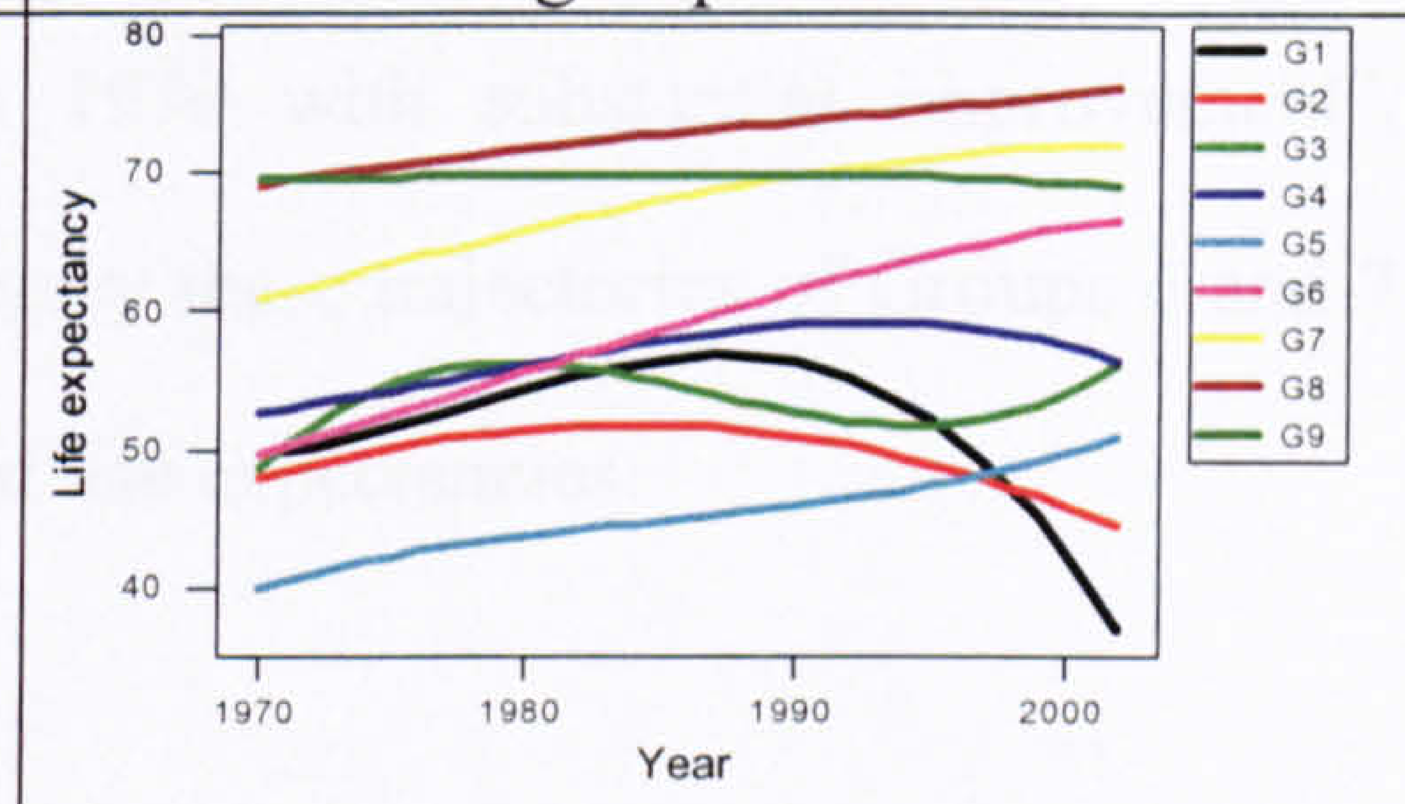
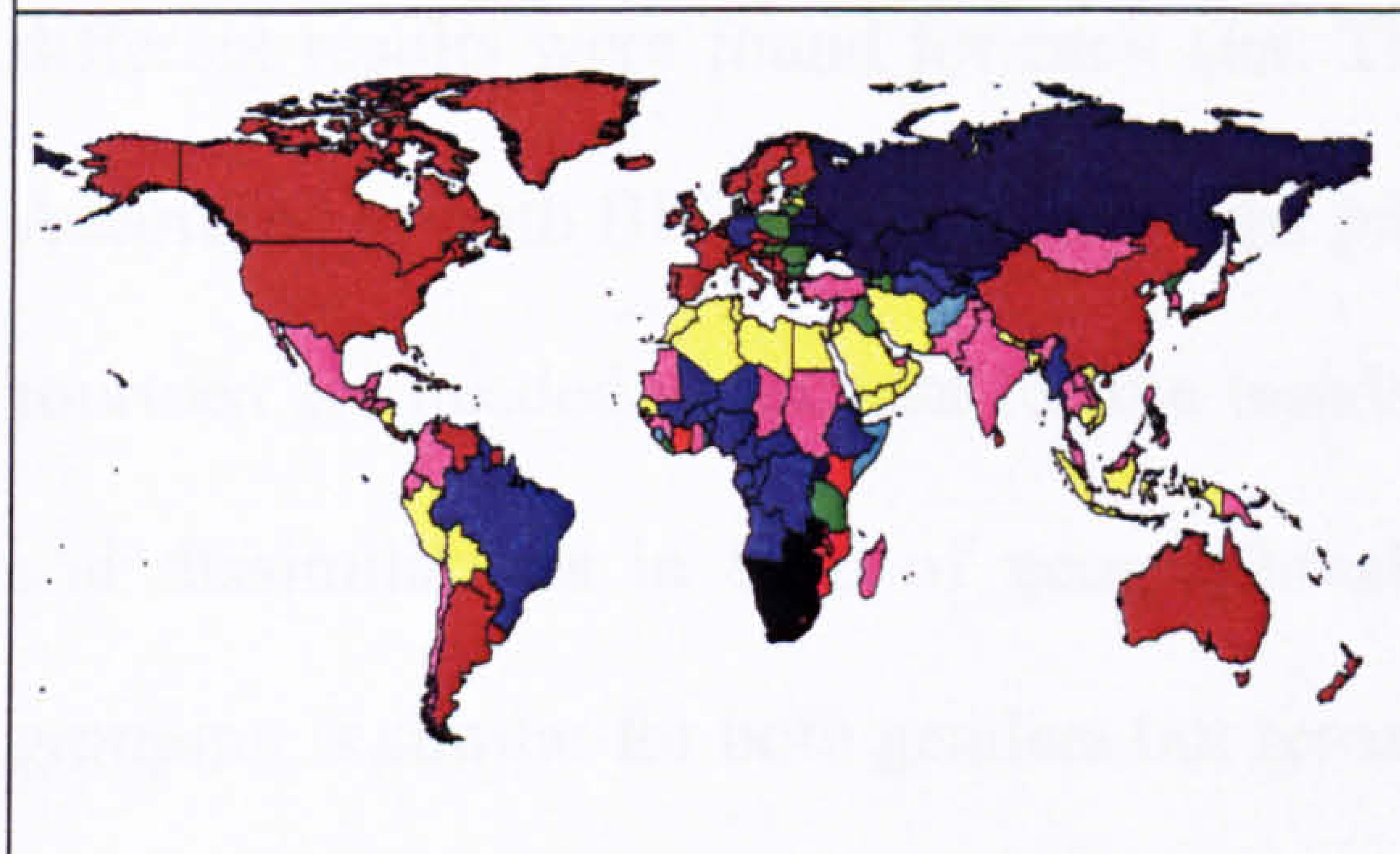
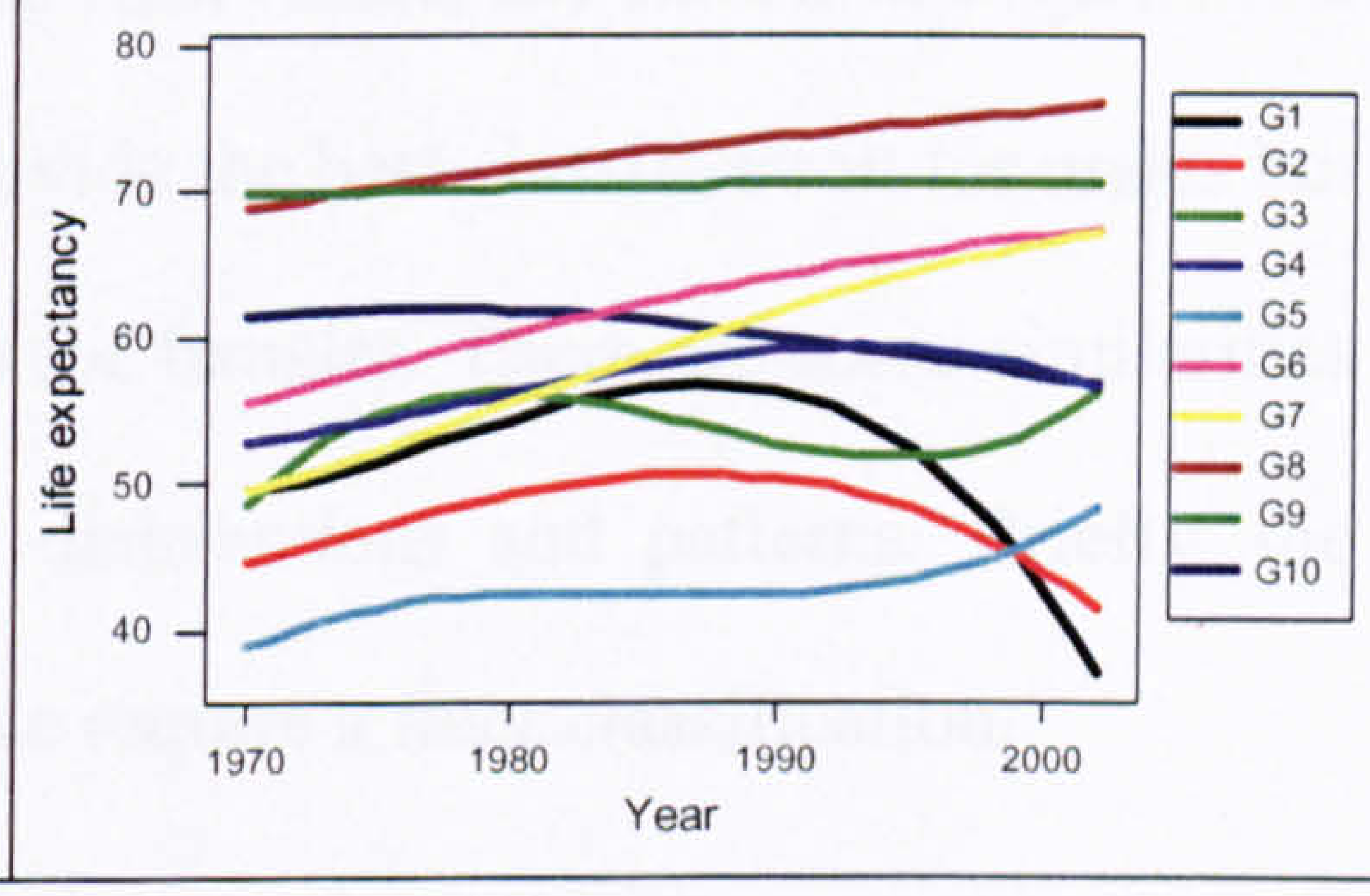
Map	**Prediction**
Model 5: five-group solution	Model 5: five-group solution
	
Model 6: six-group solution	Model 6: six-group solution
	
Model 7: seven-group solution	Model 7: seven-group solution
	
Model 8: eight-group solution	Model 8: eight-group solution
	

Figure 3.19 Trajectories for life expectancy and their distribution geographically around the world under various grouping schemes (continued)

<table>
<tr><th>Map</th><th>Prediction</th></tr>
<tr><td>Model 9: nine-group solution</td><td>Model 9: nine-group solution</td></tr>
<tr><td></td><td>
</td></tr>
<tr><td>Model 10: ten-group solution</td><td>Model 10: ten-group solution</td></tr>
<tr><td></td><td>
</td></tr>
</table>

Four distinctive, declining trajectories are Groups 1, 2, 4 and 10. Group 1 has seen the worst decline and Group 2 has also suffered a substantial shortening of life expectancy. Both of these groups are found exclusively in Sub-Saharan Africa, respectively. Group 10 consists of Russia, Ukraine, Georgia, Armenia and Kazakhstan. These are again a distinctive geographical group that has seen a decline in life expectancy from around 60 years in the 1970s. Group 4 includes most of West Africa and Brazil. This trajectory shows an initial improvement and then decline during the 1990s.

Two other similar groups are Group 6 and 7 with both showing substantial improvement. The member countries of Group 6 are quite widely distributed geographically and include the Indian sub-continent and some North African and South American countries. Group 7 includes most of North African and two South American countries (Peru and Bolivia) have been labelled "LE 55 in 1970 with substantial improvement" and "LE 50 in 1970 with substantial improvement", respectively. By the end of the twentieth century these trajectories of Groups 6 and 7 had converged to be the third highest group of life expectancies.

The analysis was then repeated for males and females separately, and somewhat different results were found for each sex. The final results are shown in Figure 3.20. According to both BICs, ten latent groups provide the best classification for males but fourteen are needed to account for the trends for females. There are some similarities and dissimilarities in term of geographical distributions and patterns. Briefly, the grouping is similar for both genders but female require a finer classification.

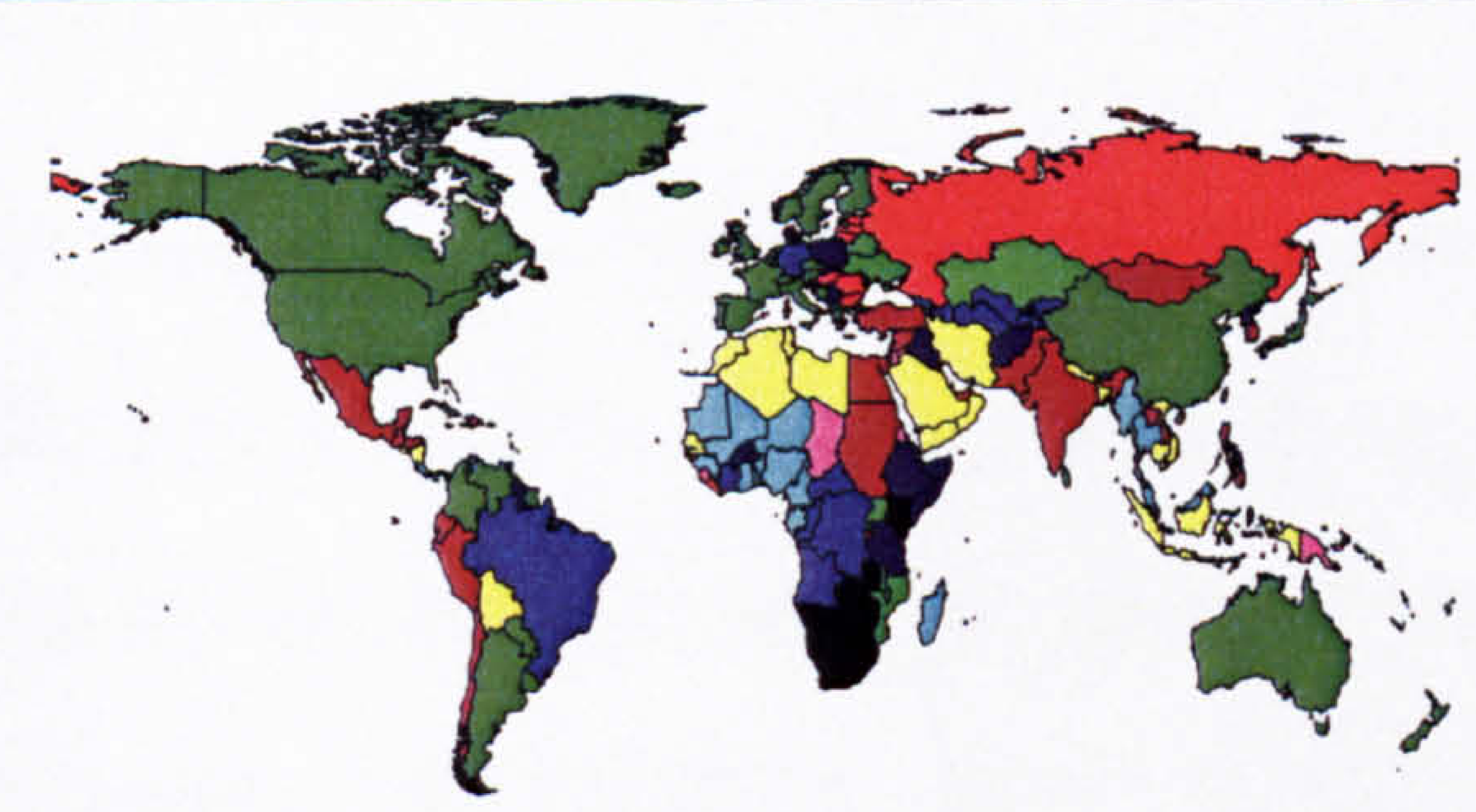
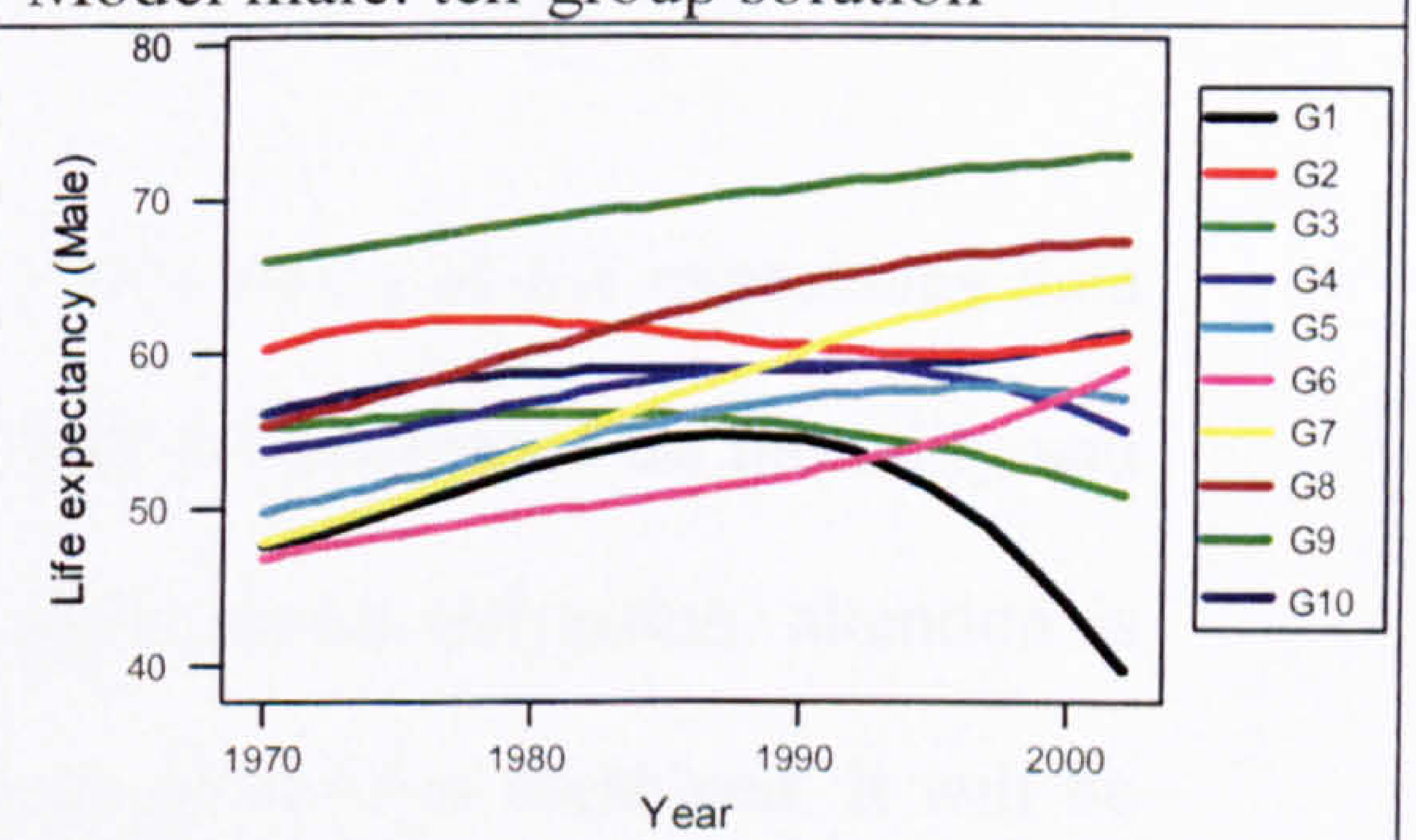
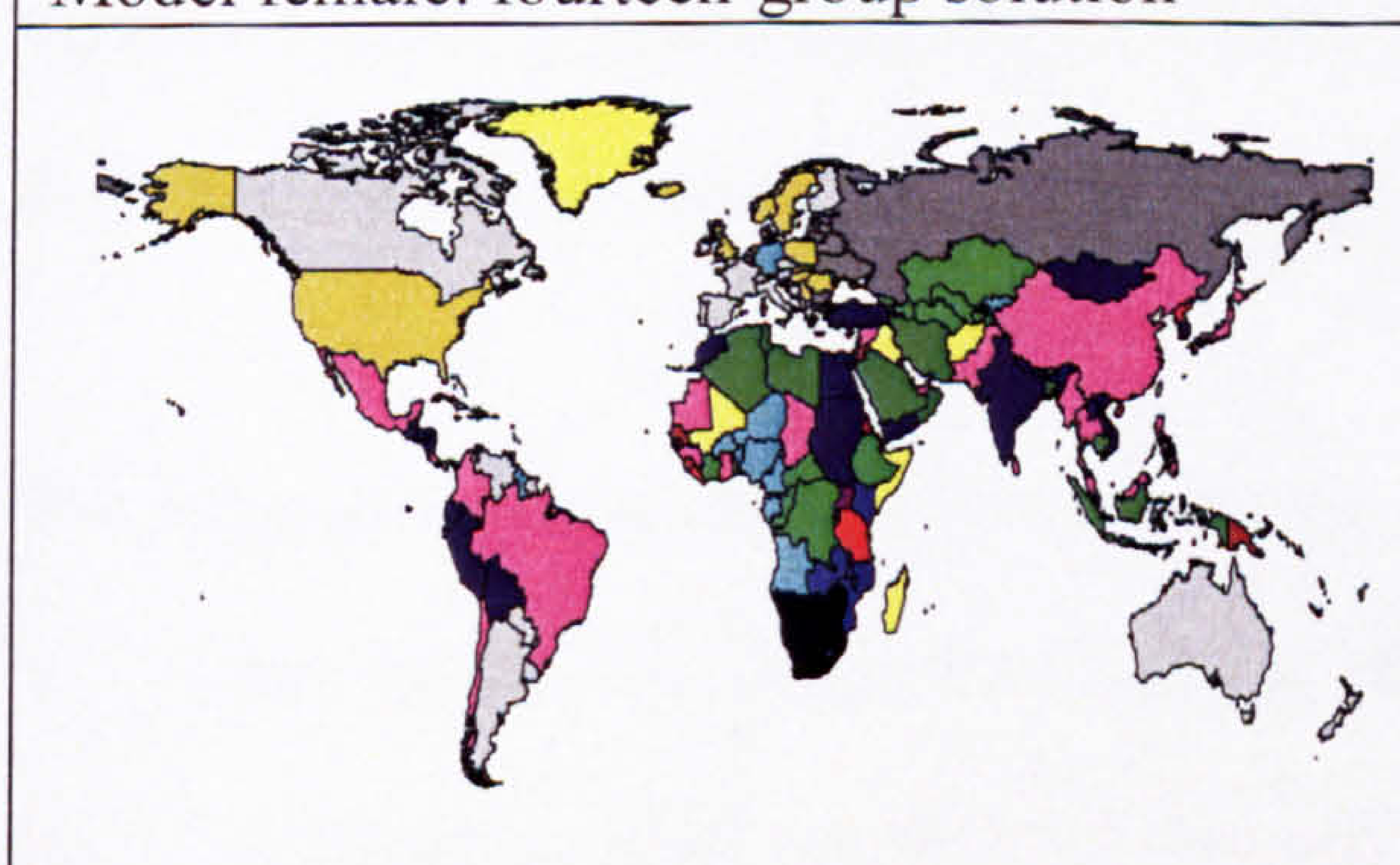
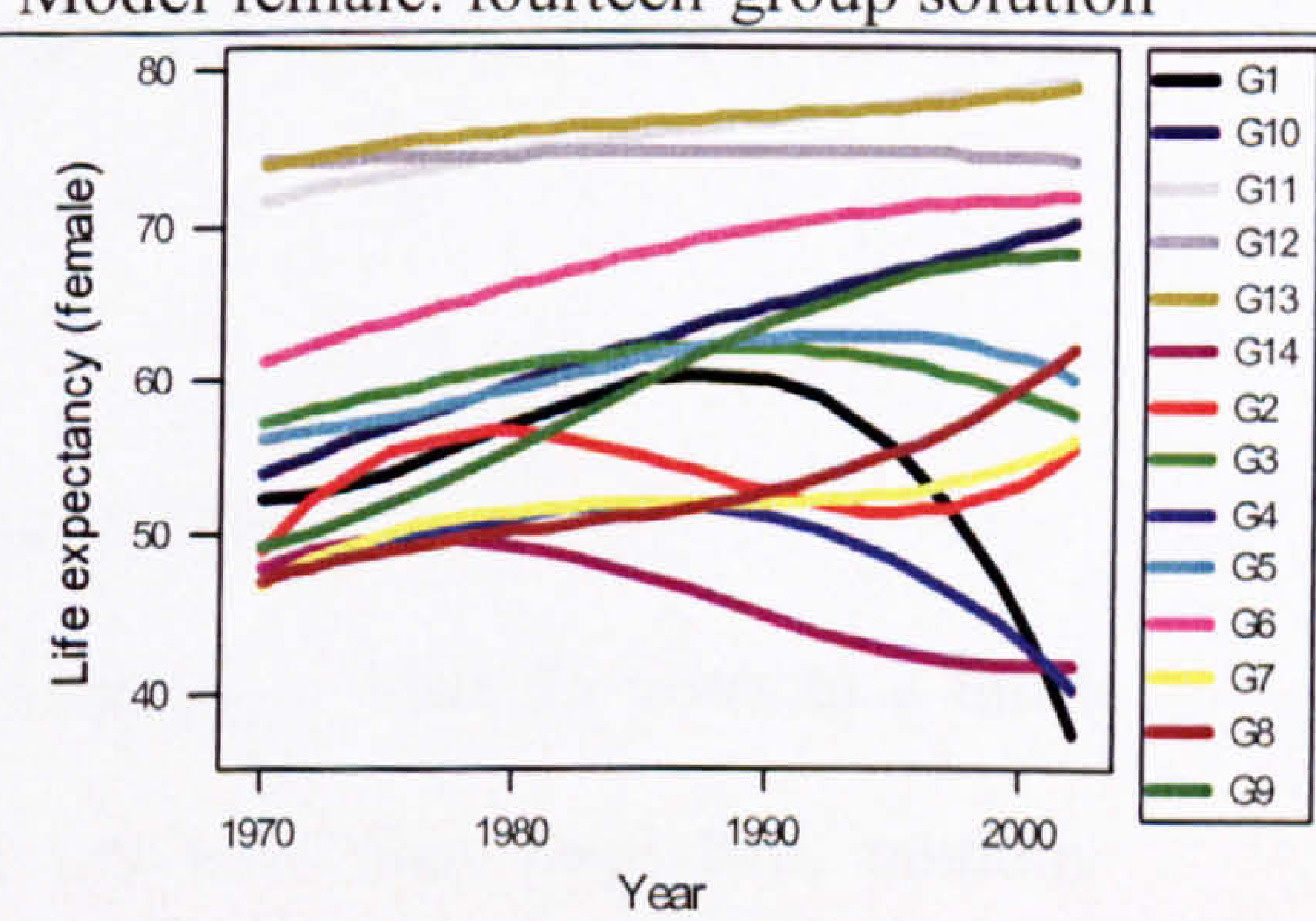
Those countries with the highest life expectancy trajectories for females which are gradually rising through time, Group 11 and Group 13, are western countries, which are very similar to the trends for males but are some 10 years higher. It is the same improving trend from around 50 to 65 years across the 30 study period for Group 7 for male and Group 9 for female with similar countries involved, (e.g. North Africa and Indonesia). The countries of India, Mongolia, Peru, Bolivia, Egypt, Turkey and Sudan, which formed Group 8 for males and Group 10 for females, have increasing longevity for both genders, but the slope is steeper for females. Sub-Sahara Africa is still the area with lowest life expectancy: it saw progress in the 1970s and 1980s but after the 1990s the decline has been dramatic.

Table 3.14 Estimates of the ten-group life expectancy model, identifying which countries belong to each group

Groups	Intercept (SD)	Year/10 (SD)	$(Year/10)^2$ (SD)	$(Year/10)^3$ (SD)	Random Intercept (SD)	Group Membership (%)	Countries includes
Group 1: LE 50 in 1970 with worst decline	**56.63** **(1.20)**	**1.45** **(0.36)**	**-5.23** **(0.15)**	**-2.11** **(0.19)**	**2.88** **(0.85)**	**3.06**	Most Sub-Saharan Africa: Zambia, Botswana, Namibia, Zimbabwe, South Africa
Group 2: 45 in 1970 with marked decline	**50.48** **(1.97)**	0.59 (0.40)	**-2.91** **(0.18)**	**-0.62** **(0.21)**	**4.32** **(1.39)**	**2.60**	Mozambique, Kenya, Ghana
Group 3: Two cycles	**54.35** **(3.62)**	**-4.44** **(0.40)**	**-0.79** **(0.17)**	**2.67** **(0.21)**	**8.00** **(2.56)**	**2.55**	Tanzania, Iraq, North Korea
Group 4: 60 in 1970 steady but decline after 90's	**58.26** **(2.11)**	**2.86** **(0.19)**	**-1.42** **(0.08)**	**-0.66** **(0.09)**	**10.61** **(1.46)**	**13.98**	Angola, Congo, Ethiopia, Gabon, Nigeria, Niger, Mali, Brazil
Group 5: Lowest LE in 1970 with improvement	**42.30** **(2.36)**	-0.26 (0.58)	**0.48** **(0.22)**	**1.28** **(0.27)**	**4.43** **(1.56)**	**2.13**	Somalia, Afghanistan
Group 6: LE 55 in 1970 with substantial improvement	**62.74** **(1.45)**	**3.86** **(0.15)**	**-0.48** **(0.06)**	-0.03 (0.07)	**9.51** **(0.10)**	**24.41**	Madagascar, Mongolia, Thailand, Laos, India, Pakistan, Sudan, Chad, Mauritania, Colombia, Mexico
Group 7: LE 50 in 1970 with substantial improvement	**59.34** **(1.33)**	**6.36** **(0.20)**	**-0.43** **(0.08)**	**-0.34** **(0.09)**	**6.37** **(0.96)**	**14.07**	Most North Africa: Morocco, Algeria, Libya, Egypt, Gambia, Oman, Yemen, Saudi Arabia, Iran, Indonesia, Peru, Bolivia
Group 8: LE of 70 in 1970 steady improvement	**72.92** **(0.45)**	**1.98** **(0.12)**	**-0.20** **(0.05)**	**0.12** **(0.06)**	**3.15** **(0.32)**	**27.83**	The USA, Western Europe, China, Australia, North Europe, Argentina, Uruguay, Paraguay, Venezuela
Group 9: LE of 70 in 1970, no change	**70.42** **(0.70)**	0.27 (0.23)	-0.10 (0.11)	-0.00 (0.12)	**2.02** **(0.49)**	**6.49**	Eastern Europe,
Group 10: LE of 62 in 1970 but subsequent decline	**60.97** **(5.06)**	**-2.06** **(0.36)**	**-0.71** **(0.18)**	0.23 (0.20)	**11.62** **(3.52)**	**2.87**	Russia, Kazakhstan, Georgia, Armenia
Level one variance: 1.31(0.02)							

For males living in the countries of the former Soviet Union, there has been a decline from about 62 years in 1970 to 58 in 2002, with Group 2 and Group 3 both showing a decline but with Group 3 typically experiencing a five year shorter life span, For females these countries do not show the declining trend with life expectancy being steady at 73 years, that is Group 12. In general Females have a wider range of life expectancies around the world than males, and this range gets wider during the study period.

Figure 3.20 Trajectories in life expectancy for males and females and their distribution geographically around the world (ten group solution)

<table>
<tr><th>Map</th><th>Prediction</th></tr>
<tr><td>Model male: ten-group solution</td><td>Model male: ten-group solution</td></tr>
<tr><td></td><td>
</td></tr>
<tr><td>Model female: fourteen-group solution</td><td>Model female: fourteen-group solution</td></tr>
<tr><td></td><td>
</td></tr>
</table>

3.4.2 Modelling group membership with time varying covariates

So far we have identified distinctive trajectories of life expectancy but we have not tried to account for these identified trends using the key variables of the Wilkinson hypothesis. I now do so by including these variables as time varying covariates. The capability of including time varying covariates in addition to time itself in the group-trajectory model has been demonstrated by Jones, Nagin and Roeder (2001) and by Jones and Nagin (2006).

The PROC TRAJ software allows the estimation of trajectories for each identified group conditioning on a set of values for time-varying covariates. That is, we can estimate different trajectories of life expectancy taking account of the changing effects of annual country income and income inequality – the key variables of the Wilkinson hypothesis. Once the model has been estimated we can also undertake a number of 'what if' scenarios.

The analysis is now limited to a 25 year (1975-1999) period of life expectancy data for the 196 countries, permitting corresponding data to be obtained for mortality and the income variables. Also, for clarity and to allow model estimation, attention is focussed on the four group solution for the males and females combined. It will be recalled that the largest reduction in the BIC reported in Table 3.11 was achieved as the number of trajectories was increased from 3 to 4.[8]

The sequence of model building and predicting is as follows.

(1) Estimating four trajectories relating life expectancy over 25 years to a third order polynomial of time in the model and with four, level two, random intercept variances.

[8] This is a common practice in the latent trajectory modelling literature.

(2) Assessing the effects of the time-varying variables income (GDP pc PPP) and income inequality derived from UTIP-UNIDO. The effects are first assessed separately and then as an interaction.

(3) Predicting changing life expectancy developmental trajectories given different underlying assumptions for income and income inequality conditions, based on the following assumptions (compared to the global average):

- all countries are as egalitarian as Sweden (Wilkinson refers to it as the most egalitarian country in the world);
- the whole World is as the most unequal country in the data set (equal to coefficient of EHII at 64) or as equal as the most equal country (equal to coefficient of EHII at 18);
- all countries are in poverty (equal to coefficient of GDP at $1,000);
- every country is affluent (as wealthy as Luxembourg);

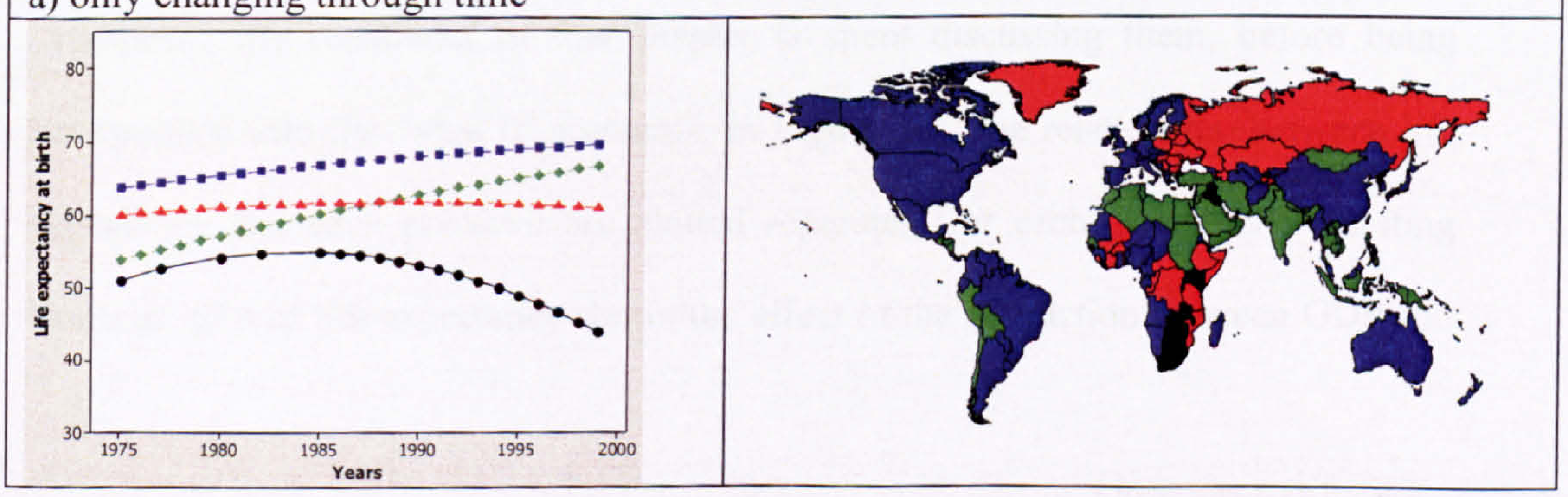
In terms of these 'what if' scenarios I have arbitrarily chosen to make the changes in the year 1986 that is roughly mid way through the period.

Figure 3.21a shows the four identified groups of life expectancy trajectories for the period 1975 to 1999. They can be labelled 'best LE (blue)', 'steady LE (red)', 'improving LE (green)' and 'worst LE (black)' There is very little change from Model 4 (which uses the slightly longer time sequence) in either the trajectory patterns or their geographical distribution (compare Model 4 in Figure 3.19). However, when taking account of time-varying country income (GDP pc ppp), it is shown to have a large influence on all of the patterns, in particular, the 'improving LE' and 'worst LE' groups – as shown in Figure 3.21b. All these estimates are plotted for when GDP is set at the global average in 1986. Countries with the improving LE trajectory (in North Africa and the Middle East) are seen to improve even more across the study period when account is taken of how rich each country is. In contrast, Sub-Sahara Africa used to have worst LE, and this becomes even worse when GDP is taken into account. Giving changing GDP, the gap between 'best LE' (most western countries) and 'steady LE' (former Soviet Union States) closed in the late 1990s. Indeed given how rich it is, the USA has 'swapped' group membership from 'best LE' into 'steady LE'.

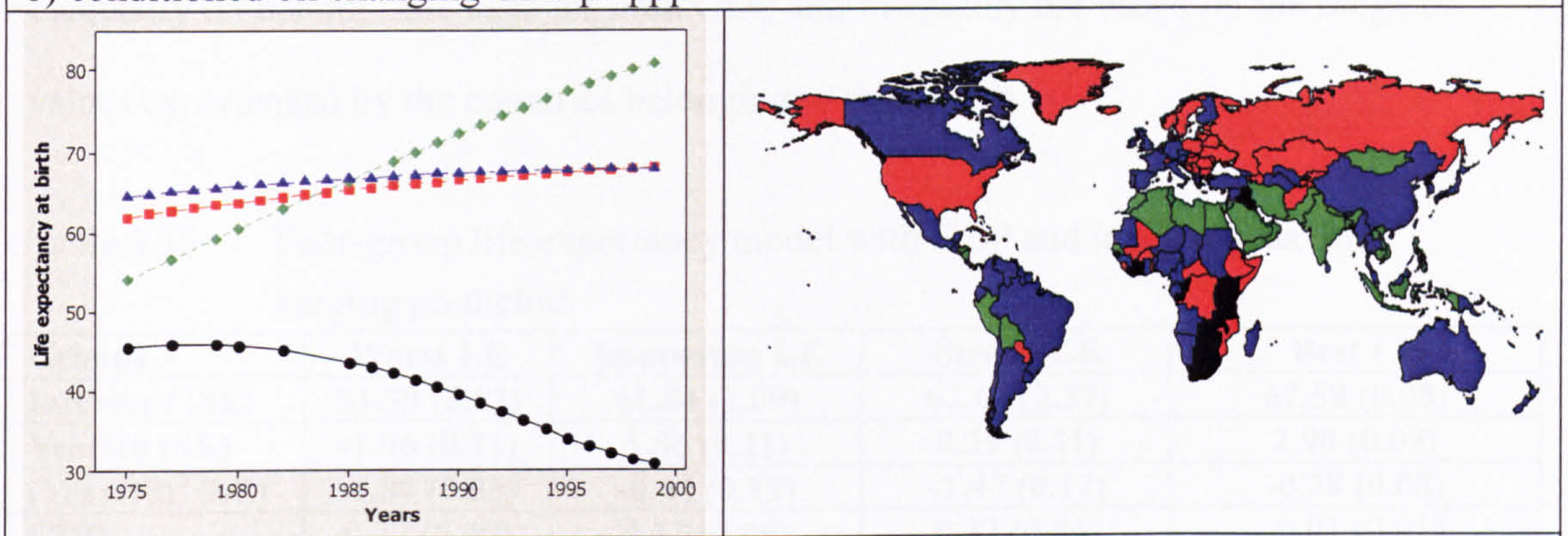
In Figure 3.21c, time-varying income inequality is included in the model as the sole time-varying term. The effect of income inequality does not change the patterns a great deal. Indeed, the 'slopes' of the trajectories are markedly the same. The only real change is that in taking account of income inequality, the 'improving LE' group now has a consistently higher life expectancy than the 'steady LE' one. Again, these predictions are made for when inequality it at its average value in 1986. The detailed information on the estimates of the model with GDP and inequality (EHII) as main effects and also as interaction terms are given in Table 3.14; the trends are plotted in Figure 3.21d.

Figure 3.21 Life expectancy trajectories and their geographical distribution around the world under four-groups solution before and after including income and inequality time varying covariates

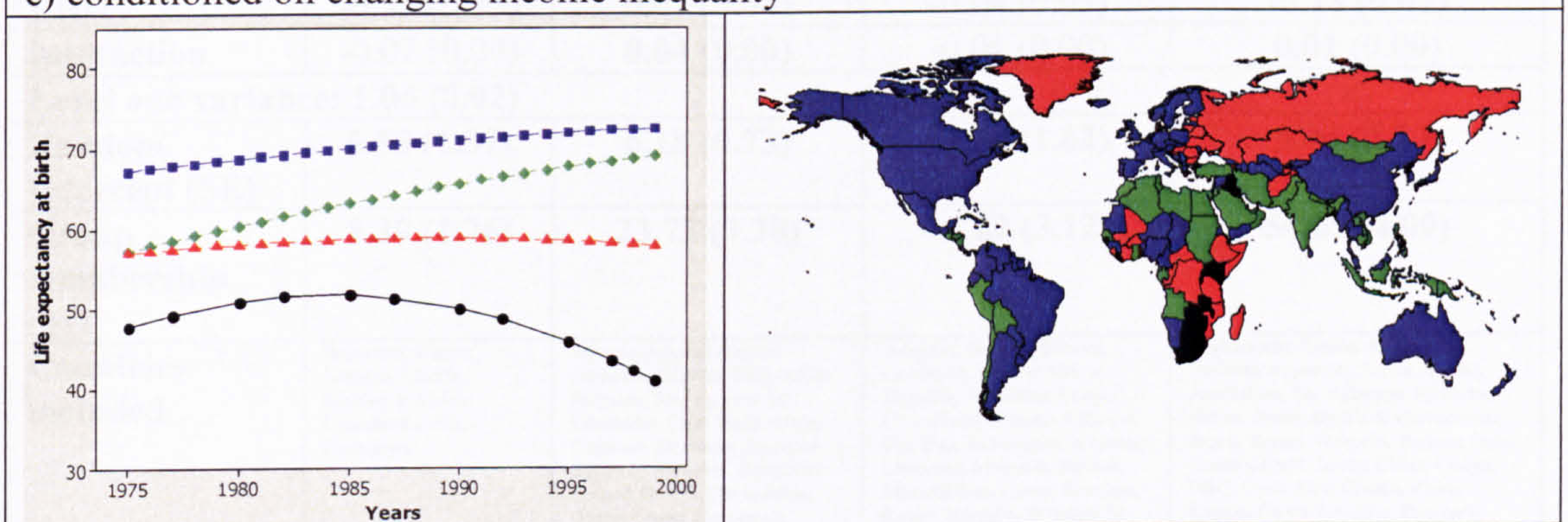
a) only changing through time



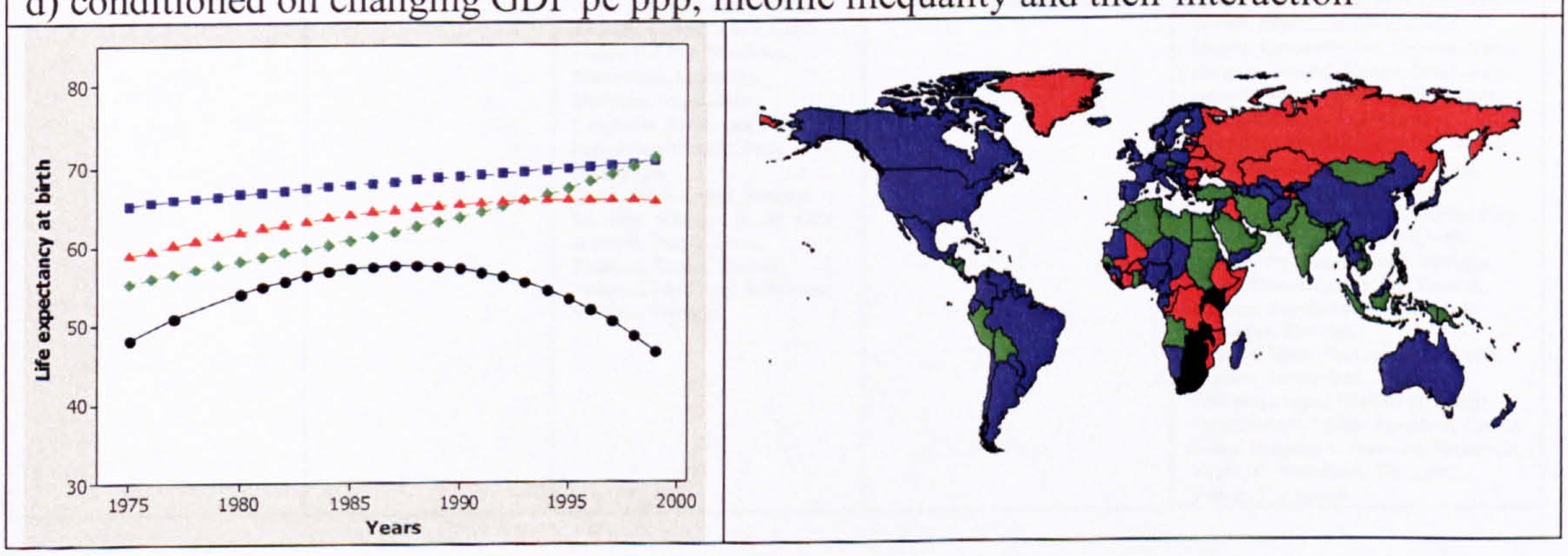
b) conditioned on changing GDP pc ppp



c) conditioned on changing income inequality



d) conditioned on changing GDP pc ppp, income inequality and their interaction



Given the central importance of these results as a first test of the Wilkinson hypothesis, the remainder of this chapter is spent discussing them, before being incorporated into the 'what if' scenarios. In Figure 3.22 the relationship between life expectancy and each predictor are plotted separately for each group. The resulting 'contour' plot of life expectancy shows the effect of the interaction between GDP and inequality on health; the axes for both GDP and inequality are based on the range of values experienced by the countries belonging to that group.

Table 3.15 Four-group life expectancy model with GDP and inequality as time varying predictors

Groups	Worst LE	Improving LE	Steady LE	Best LE
Intercept (SE)	53.30 (3.82)	61.84 (1.09)	62.67 (2.39)	67.58 (0.98)
Year/10 (SE)	-1.96 (0.21)	5.56 (0.11)	0.35 (0.11)	2.90 (0.08)
$(Year/10)^2$ (SE)	-4.80 (0.25)	-0.44 (0.12)	-1.47 (0.17)	-0.28 (0.08)
GDP	0.32 (0.40)	0.13 (0.06)	0.33 (0.04)	-0.01 (0.01)
EIHI	-0.37 (0.35)	0.07 (0.04)	-0.04 (0.05)	-0.18 (0.02)
Interaction	-0.07 (0.04)	0.04 (0.00)	-0.01 (0.00)	0.01 (0.00)
Level one variance: 1.06 (0.02)				
Random Intercept (SE)	5.55 (1.37)	6.18 (0.73)	11.58 (1.62)	9.00 (0.70)
Group membership (%)	5.39 (1.76)	23.73 (3.38)	16.22 (3.12)	54.65 (4.09)
Countries included	Botswana, Kenya, Lesotho, Liberia, Malawi, S Africa, Uganda, Zambia, Zimbabwe	Albania, Algeria, Angola, Australia, Bahrain, Bangladesh, Belgium, Bhutan, Bolivia, Cambodia, Cape Verde, Chile, Comoros, Dominica, Ecuador, Egypt, El Salvador, Equatorial, Guinea, Gabon, The Gambia, Ghana, Guam, Guatemala, Honduras, Hong Kong (China), India, Indonesia, Iran, Jordan, Kiribati, Kuwait, Laos, Libya Macao (China), Maldives, Micronesia, Mongolia, Morocco, Nepal, New Caledonia, Nicaragua, Pakistan, Papua New Guinea, Peru, Philippines Qatar, Saudi Arabia, Senegal, Slovenia, Solomon Is., St. Kitts & Nevis, Sudan, Syria, Thailand, Tonga, Tunisia, Turkey, United Arab Emirates, Vanuatu, Vietnam	Armenia, Belarus, Burundi, Cameroon, Central African Republic, Colombia, Congo, Cote d'Ivory, Estonia, Ethiopia, Fiji, Iraq, Kazakhstan, N Korea, Lithuania, Malaysia, Mexico, Mozambique, Oman, Romania, Russia, Rwanda, St. Lucia, St. Vincent & the Grenadines, Swaziland, Tanzania, Ukraine	Afghanistan, Samoa, Antigua & Barbuda, Argentina, Aruba, Austria, Azerbaijan, The Bahamas, Barbados, Belize, Benin, Bosnia & Herzegovina, Brazil, Brunei, Bulgaria, Burkina Faso, Canada, Chad, Jersey, China, Congo, DRC, Costa Rica, Croatia, Cuba, Cyprus, Czech Republic, Denmark, Djibouti, Eritrea, Finland, France, French Polynesia, Georgia, Germany, Greece, Greenland, Guadeloupe, Guinea, Guinea-Bissau, Guyana, Haiti, Hungary, Iceland, Ireland, Israel, Italy, Jamaica, Japan, S Korea, Kyrgyzstan, Latvia, Lebanon, Luxembourg, Macedonia, Madagascar, Mali, Malta, Martinique, Mauritania, Mauritius, Moldova, Myanmar, Namibia, Netherlands, Netherlands Antilles, New Zealand, Niger, Nigeria, Norway, Panama, Paraguay, Poland, Portugal, Puerto Rico, Reunion, Sao Tome & Principe, Seychelles, Sierra Leone, Singapore, Slovakia Somalia, Spain, Sri Lanka, Suriname, Sweden, Switzerland, Taiwan, Tajikistan, Togo, Trinidad & Tobago, Turkmenistan, United Kingdom, United States, Uruguay, Uzbekistan, Venezuela, Virgin Is., West Bank, Gaza Strip, Yemen, Yugoslavia

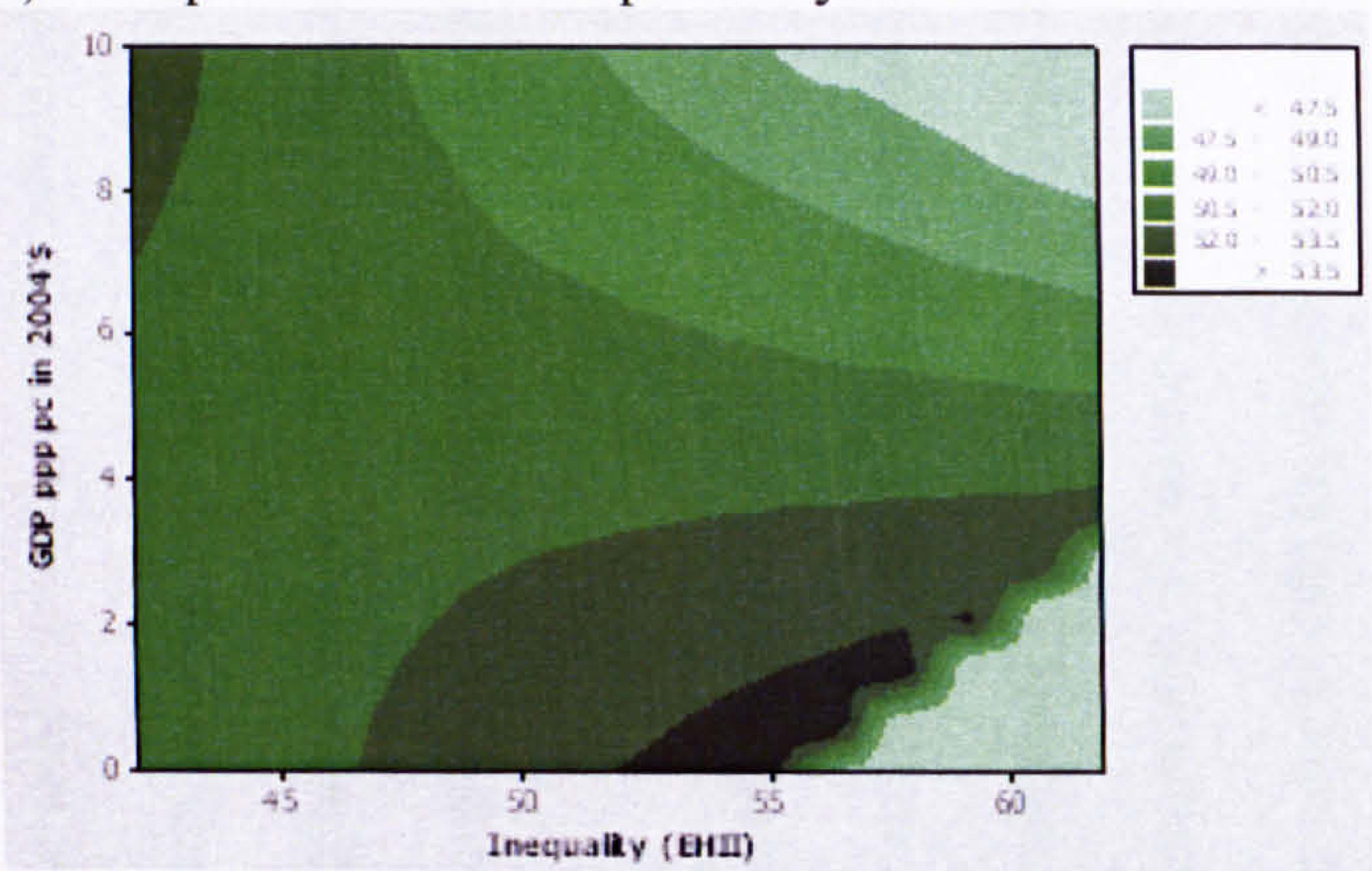
Taking each group at a time, Group 1 is found to have the worst life expectancy and represents Sub-Saharan Africa. For this group, with its low values of GDP and high values of inequality, the contour plot (Figure 3.22a) shows that higher life expectancy is to be found where the country is relatively rich (within the group) and where the country is relatively equal (in terms of income). The estimated parameters are the largest in comparison to the other groups, however none of the three terms (main effects and interaction) are significant – this may reflect that with a relatively small number of countries there is insufficient power to detect significant effects.

The next contour plot (Figure 3.22b) shows the results for the improving LE Group and includes countries in North Africa and the Indian subcontinent. Here the main effect for GDP is significant and positive, as is the interaction; the effect for inequality is positive but not significant. Greatest life expectancy is to be found in the richer countries and the ones that have greater income inequality. Group 3, which contains the countries of the former Soviet Union and central Africa, is shown by Figure 3.21d to have a steady increase in life expectancy when account is taken of GDP and inequality with regard to the estimates of Table 3.14, while there is a significant positive effect for GDP, the effects for the inequality and the interaction are not significant. This is shown clearly on the contour plot of Figure 3.22c where life expectancy only increases as GDP increases. In this group of countries, inequality has little effect.

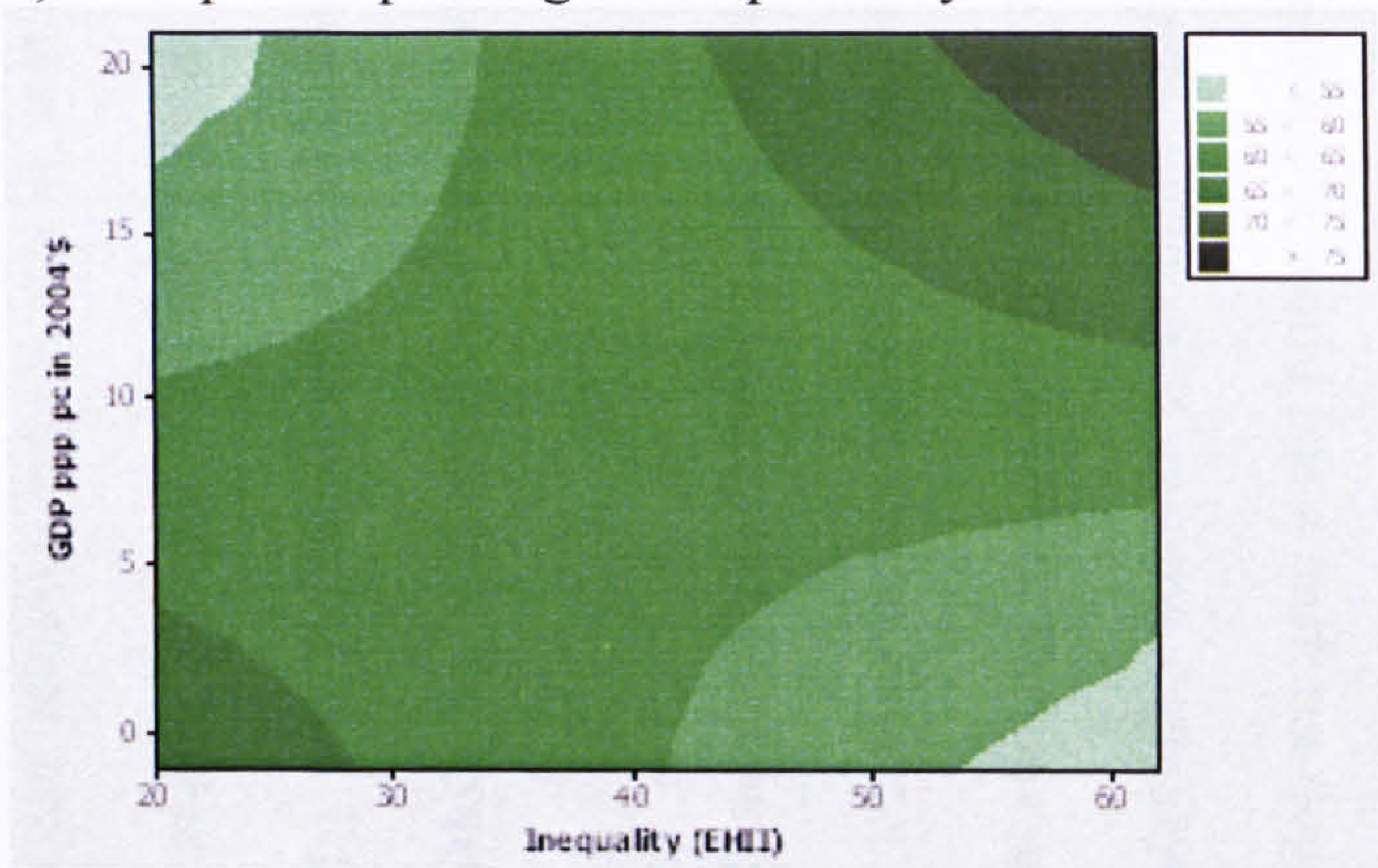
Finally, we come to the largest group of countries, with the best life expectancy and which represents the advanced economies of the West. As correctly predicted by the Wilkinson hypothesis, the effect for GDP is small and insignificant; however and again as correctly predicted by the hypothesis, the effect for inequality is significant and negative. The contour plot shows that the highest life expectancy is for the lowest inequality.

Figure 3.22 Visualising the effects of GDP and inequality in life expectancy

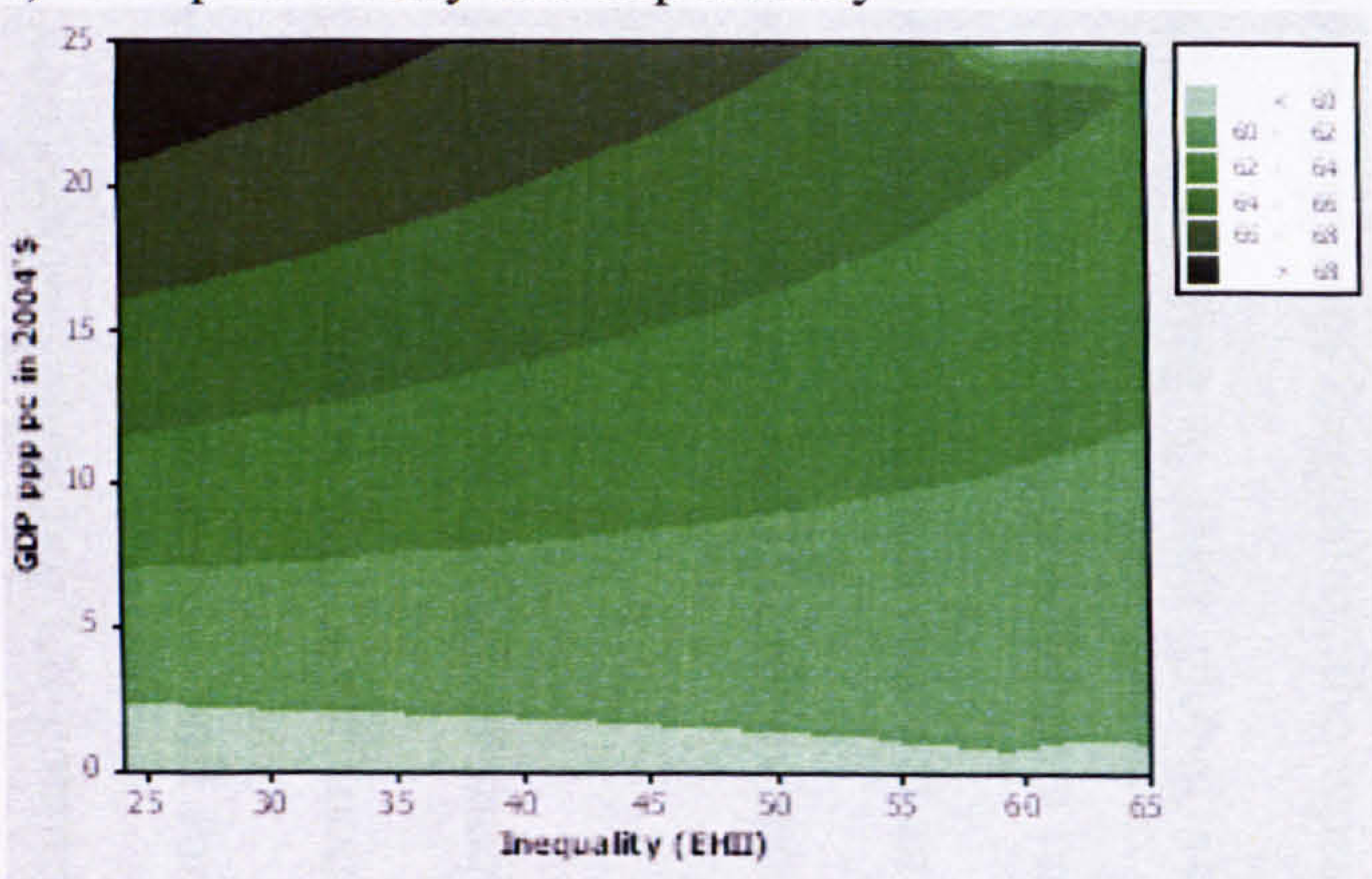
a) Group 1: Worst life expectancy



b) Group 2: Improving life expectancy



c) Group 3: Steady life expectancy



d) Group 4: Best life expectancy

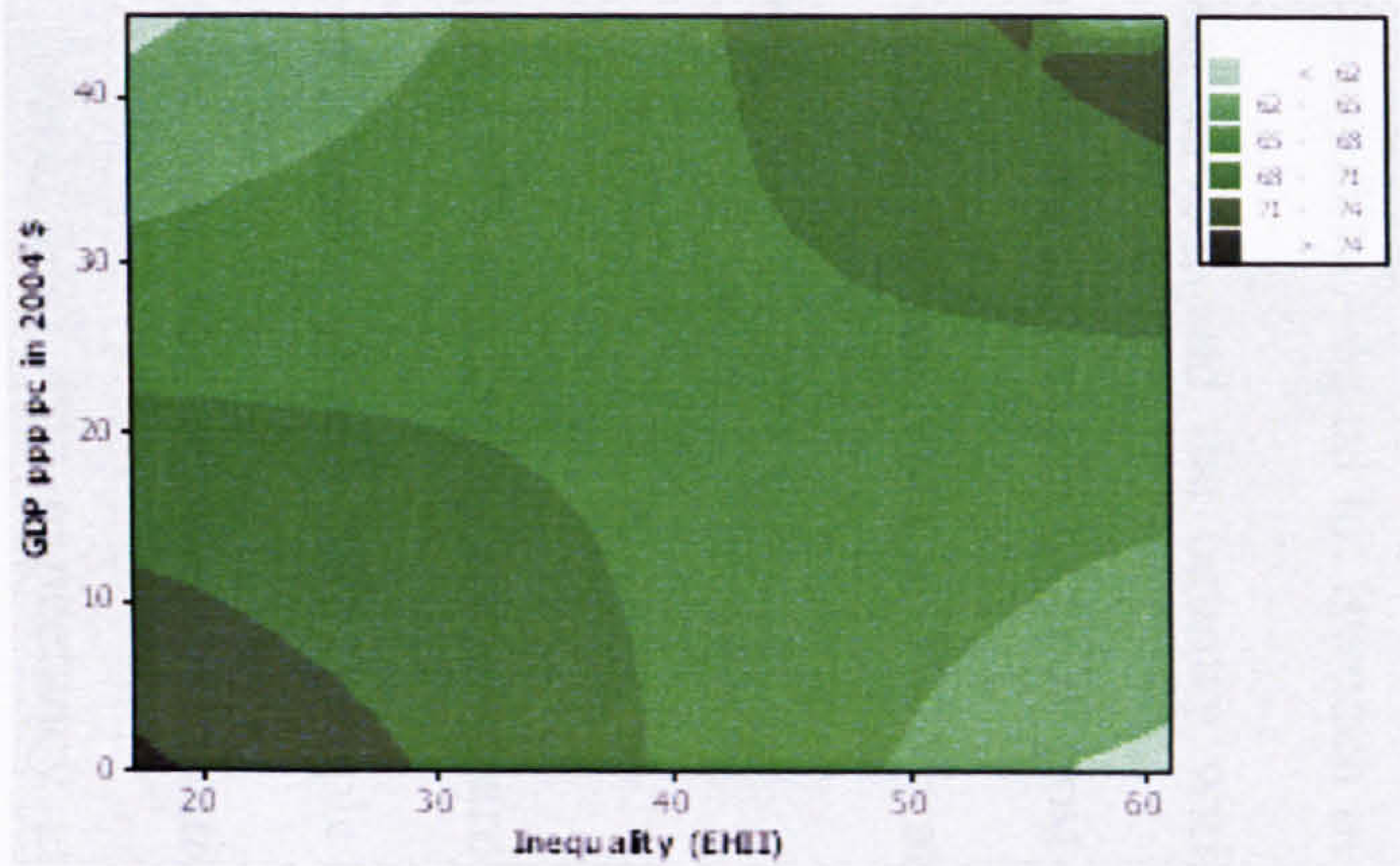


Figure 3.23 shows the effect of different 'what if' scenarios on life expectancy trend of each group trajectory. Before 1986 the values of income and inequality were kept at their global average value (GDP of ten thousand US dollars and an inequality of 40), but after 1986 a range of scenarios are applied. These are mentioned above and assume that the world has a uniform degree of inequality – equal to: Sweden in 1981; the country with most income equality (Latvia, 1975); and the country with least income equality (Togo, 1999). Various assumptions are also made of GDP post-1986: the same as the poorest in the data set (Bosnia and Herzegovina in 1975); and the same as the richest country (Luxembourg in 1999).

The plots of the countries with the worst changing life expectancy patterns are given in Figure 3.23a. Both inequality and GDP have sizeable effects with people living longer when a country' income higher and with narrower income distribution. In comparison to the global average inequality (black line), life expectancy at birth increases by about 8 and 4 years when every country become the most equal (red line) and as equal as Sweden, respectively (blue line), but it decreases by 8 years if they are the most unequal countries (green line, the higher degree of inequality, the worsening in life expectancy). However, the effect in GDP comparing to their global average income (black line) is that the life expectancy rises by about 10 years when a country is more developed with highest income as Luxembourg but declines more than three when a country is in most disadvantage material life as Bosnia and Herzegovina.

In countries of improving life expectancy in Figure 3.23 b), the effect of inequality not only has smaller differential impact between most equality and most inequality but also with opposite order, in that more unequal income distribution will improve life expectancy. Similar income effect pattern happens to improving LE countries as countries with worst LE but with smaller differential degree. There is no significant inequality effect but large income effect to countries with steady LE. Income is the most important factor for those Former Soviet Unite countries to improve their life expectancy but inequality does not matter much (see Figure 3.23 c).

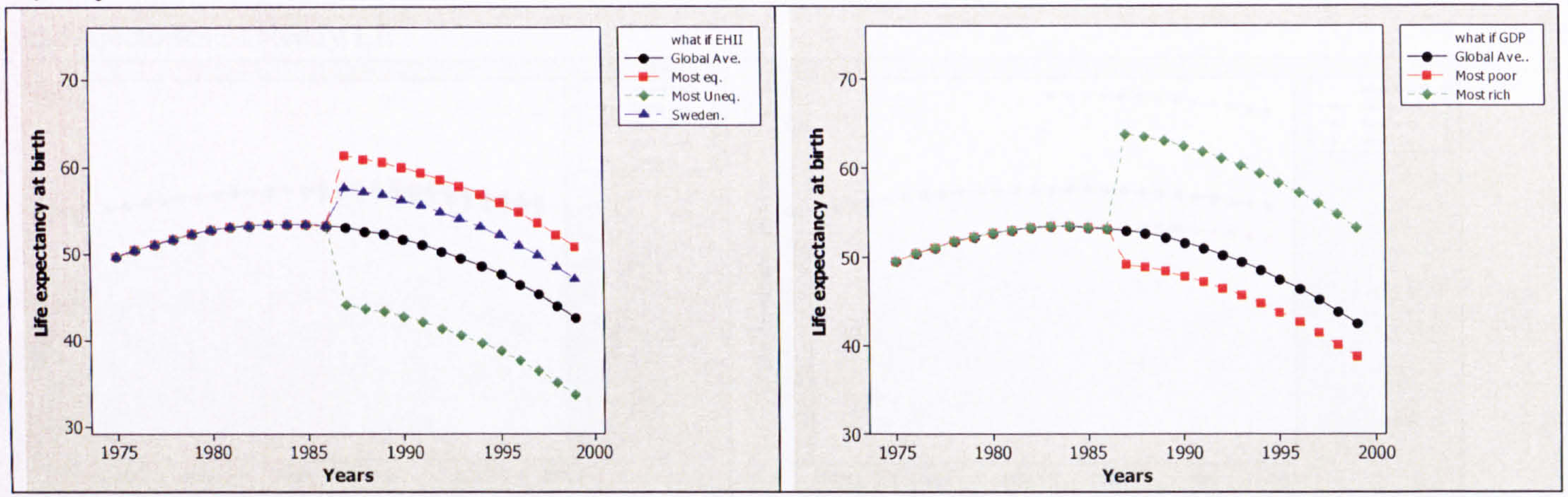
In Figure 3.23 d), the patterns for Best LE countries are slightly increasing for both effects of income and income inequality on global average. If a country is as equal as Sweden or the most equal country, life expectancy increases by less than 5 years and the later is increasing even more than the former. However, the degree of a country's income has no significant impact in life expectancy. This supports Wilkinson's hypothesis that what matters to a developed country is the degree of income distribution but not absolute income.

3.5 CONCLUSIONS

Over the last three decades, the global life expectancy of both males and females has increased, but the trend at the millennium has levelled off. Women having consistently have higher life expectancy than men, both sexes have experienced this recent levelling. There is a relatively large between-country variation that converged towards the mid 1980s but diverged subsequently, with the variation being greater for women than for men. When the group trajectory model is applied on the dataset compiled for this research, it has been able clearly to identify groups of countries with distinctive trends: some groups with increasing life expectancy trajectories; some with declining life expectancies; and some showing increase and decrease during the study period.

Figure 3.23 The changing trajectories of life expectancy based on different assumptions of country's income and inequality after 1986

a) Trajectories of 'worst LE'



b) Trajectories of improving LE

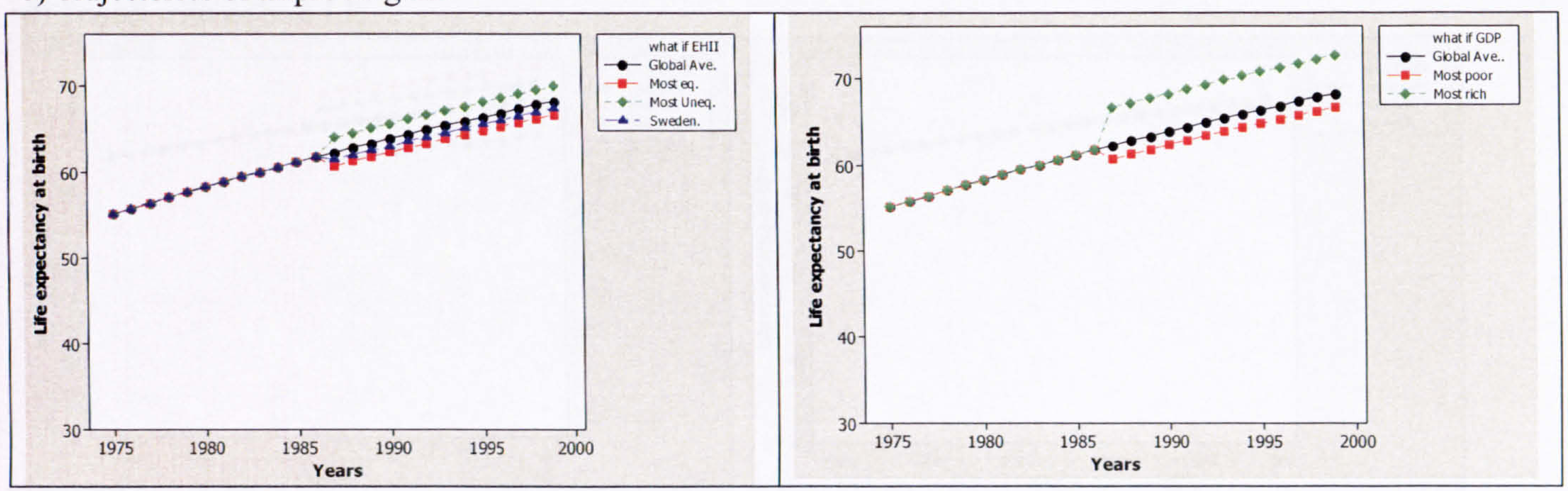
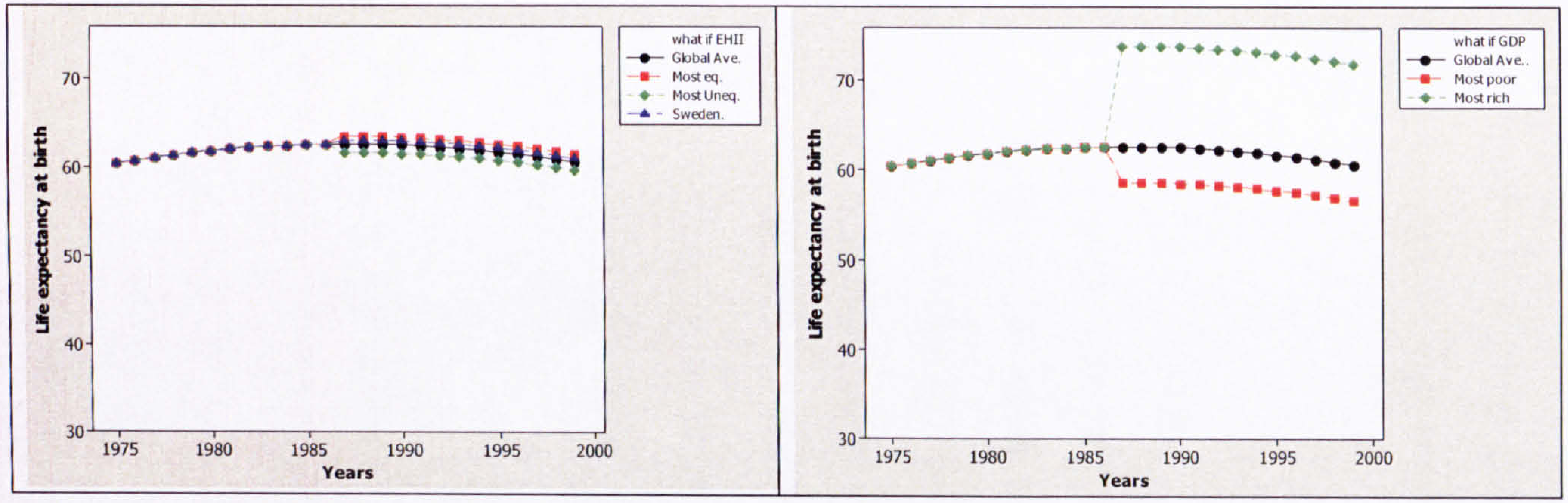
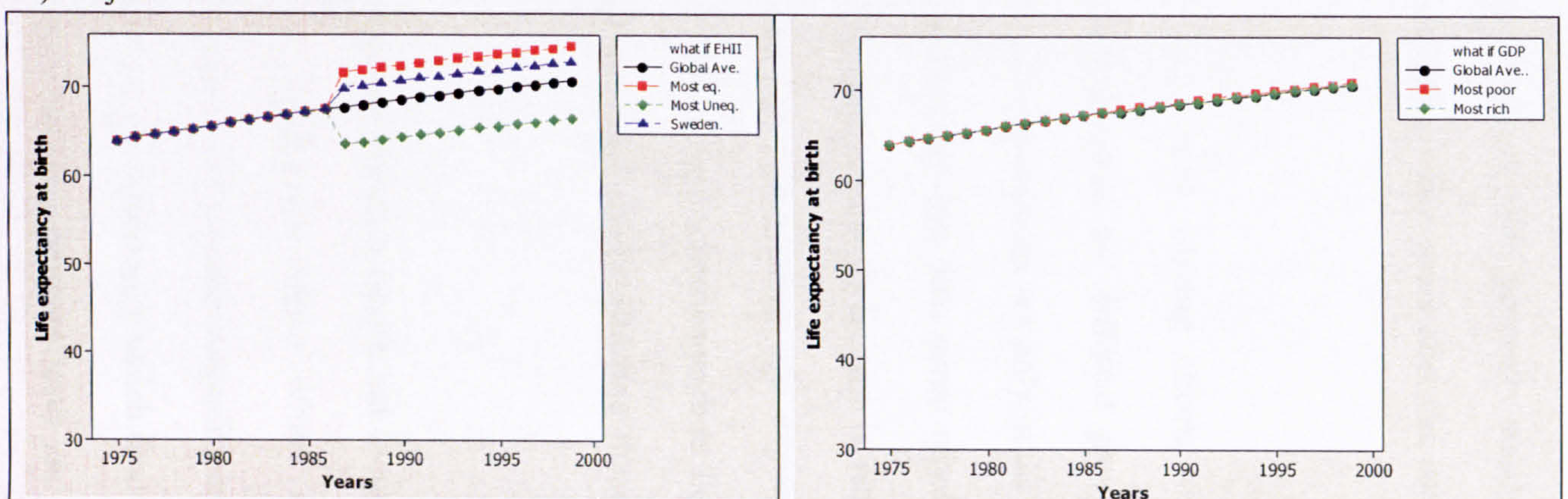


Figure 3.23 The changing trajectories of life expectancy based on different assumptions of country's income and inequality after 1986 -- continued

c) Trajectories of steady LE



d) Trajectories of best LE



These trajectories have been shown to have a clear spatial expression. Importantly, the results from this study suggest it is important and necessary to look at developmental trajectories of countries in a number of small groups, because the effect on life expectancy of a country's income and income inequality is different from one group to another. For example, and at the extremes, western countries have the highest life expectancy which increases through time, whereas Sub Sahara Africa countries have high mortality that has substantially worsened during the last thirty years. Former Soviet bloc countries also show a very distinctive pattern, with generally steady and high life expectancy during the study period but losing some years after the collapse of communism.

GDP and income inequality have been shown to have varying effects on life expectancy – and importantly this effect is differentiated for different groups of countries. Support has been found for the Wilkinson's hypothesis not only in the most wealthy economies (which Wilkinson generally studies) but also some other less wealthy countries. In this research it also appears that a country's income is vital for countries (mostly North Africa countries) with large increasing life expectancy positively related to GDP, as well as steady life expectancy countries such as Former Communist States. However, the degree of income distribution affecting these two groups of countries is insignificant.

As expected, increases in GDP are found to improve people's health but inequality worsens it in developing countries such as Sub-Sahara Africa which suffer disadvantaged living standards. However, both income and income inequality are not statistically significant. In the countries with the best life-expectancy which tend to be developed countries, when analysis is undertaken at this aggregate level and over time; the trends are as predicted by Wilkinson. GDP has little effect on longevity while in countries with unequal income distributions, people live relatively shorter lives.

Chapter 4

Global variations in self-reported health: an analysis of the World Values Survey

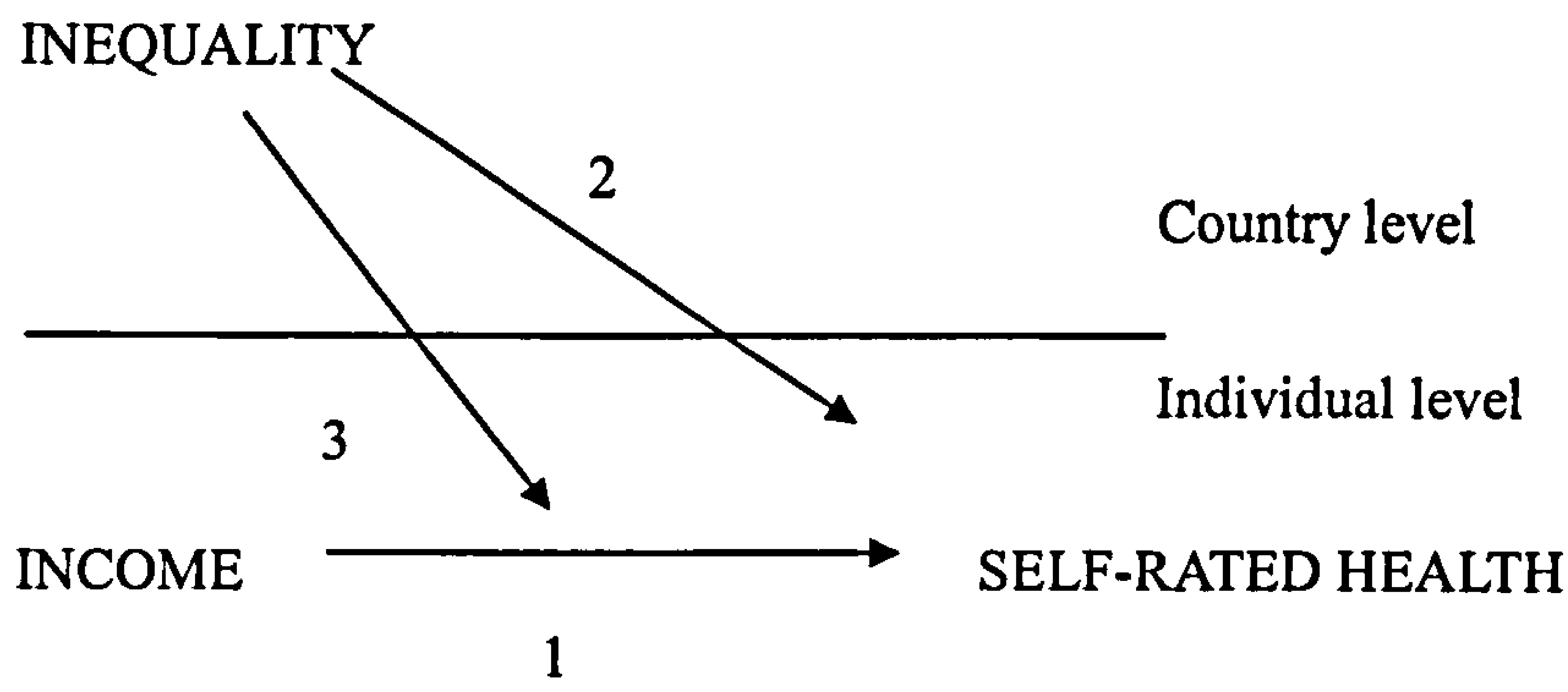
4.1 INTRODUCTION

As discussed in Chapter Two, Wilkinson relies to a large extent of empirical support for his relative-income hypothesis on global studies of income inequality and mortality where 'societies' are defined operationally as countries. However, as demonstrated in Chapter Two, Wilkinson's aggregate analysis is potentially flawed for when there is a non-linear relationship between income and mortality; a spurious aggregate relation will be found between mean mortality and the within-country variability (that is inequality) in income. A study is therefore needed whereby an appropriate measure of health is related at the individual level to income and simultaneously to average income and income inequality. Given that Wilkinson developed his hypothesis in relation to countries as societies, it is crucial to examine a wide variety of countries with a wide range of average income and inequalities. Moreover, as the Wilkinson argument implies that once countries reach a certain stage of development (which he operationalizes as GNP $5000 per capita in 1990), inequality takes over from per capita income as the primary determinant of health, it is important that the hypothesis is examined over time as countries become more developed.

There are two important objectives in this chapter. First, I am going to analyse how people's health outcome is related to income at an individual level after taking account of individual demographic characteristics such as age, sex and marital status. This is to determine whether the first element of Gravelle's (1998) critique of Wilkinson's hypothesis is supported, in particular, if there is a non-linear relationship between income and mortality. Second, I will then examine whether individual ill health is related to a country's wealth and inequality once individual income is taken into account to see if, as Gravelle claims, the relative income inequality hypothesis is an artefact of aggregate data analysis.

The argument was made in Chapter Two that no individual-based studies to date have evaluated the hypothesis at the scale they were originally formulated, that is countries. This is remedied in this chapter by using individual data derived from the World Values Survey (WVS) analysed by an appropriate multilevel modelling methodology. The data set has been assembled from 69 country cohorts across 4 waves: 1981, 1990, 1995-97 and 1999-2001. The WVS not only includes a variety of variables about individual health and feelings such as self-rated health and happiness etc, but also has information on personal demographic characteristics and individual income. Consequently, the main question to be addressed using the WVS data is whether income at the individual level and income inequality at the national level affects the self-rated health of individuals.

In addition, the hypothesis will be tested whether the relationship between income and self-rated health is stronger in more equal countries than unequal ones. This is a cross-level interaction between individual and country characteristics. It is summarized in a graphical way below so that path 1 is the individual income effect, path 2 is the macro, country-level inequality effect and path 3 is the cross-level interaction whereby, for example, the difference between poor and rich individuals is exacerbated in countries with unequal income.



* Micro (=1), Macro (=2), and cross-level interaction (=3) effects on Self-rated health.

The specific research goals for this chapter are

- To assess the between-country and between-wave differences for self-rated health
- To evaluate the individual effect of demographic characteristics and socio-economic indicators on self-rated health
- To examine the form of the relation between individual income and health to assesses if it is linear or non-linear
- To examine the relationship between country income and self-rated health for countries above and below five thousands US dollars in GDP (purchase power parity) per capita in 1990 separately
- To estimate the relationship between self-rated health and income inequality for countries above and below five thousands US dollars in GDP (purchase power parity) per capita in 1990 separately
- To estimate whether the individual income effects on self-rated health are influenced by the income inequality of the countries
- To consider the remaining between-country variation after taking account of individual income, GDP, and inequality

- To evaluate the hypothesis that there are unhealthy communities (defined as countries) as well as unhealthy individuals.

The remainder of the chapter is in three parts; data, methodology, and results with discussion. First, in the data section, the question employed to operationalize ill-health will be listed and how the WVS dataset needs to be manipulated to achieve the objectives outlined above will be considered. Second, the technique of multinomial multilevel modelling will be described in detail. This is used here because the response variable has three categories (good, fair and poor health) thereby requiring multinomial modelling; and the WVS data has a four–level structure of outcome measurement nested within individuals who are nested within different waves, with countries at the top of the hierarchical structure, thereby requiring multilevel modelling. This methodology allows us to model the micro level (age, sex and individual income), the macro level (mean income and income inequality) and the cross-level interactions between individual income and income inequality. The final part of the chapter reports a series of models of increasing complexity to evaluate the Wilkinson's hypothesis and a number of alternatives.

4.2 DATA

4.2.1 The World Value Survey data

In this section, the main focus is on what data are required to test the Wilkinson's hypothesis and how the World Value Survey (WVS), which provides individual records, meets these requirements. There are also brief discussion of two country-level data sets, representing a country's income and its income distribution. These two latter data sets have been already considered and employed in Chapter Three. They provide information across the study period as well as for a large range of countries. Preparing an appropriate global dataset involved a considerable amount of data processing including collation, linkage and cleaning. The undertaken process will be outlined here briefly.

Wilkinson in his work has mainly used mortality statistics to evaluate his hypothesis. Indeed, mortality statistics are widely used, and are the most specific indicator for monitoring the health of a general population (Alderson, 1988), especially when considering health variations across different nations (Shaw *et al.,* 2002, p87). However, death records are usually only available in aggregated form, and such data do normally include linked key variables such as income; also routine data usually have very little information on those that form the denominator, individuals who are still alive.

Therefore an alternative indicator of health is needed and I have used the "self-rated health" as this is measured at the individual level in the cross-sectional cohort study of the WVS data set. This variable has been collected by asking respondents "All in all, how would you describe your state of health these days? Would you say it is excellent, very good, good, fair, poor or don't know?". There is an extensive amount of literature (reviewed in Idler and Benyamini, 1997) that shows this indicator to be a valid instrument, and capable of assessing mortality risk independent of other medical, behavioural and psychosocial risk factors. Furthermore, it has also been shown in longitudinal studies that self-rated health predicts the onset of disability (Wilcox et al, 1996). Normally, this question is asked conditionally on a person's age, that is with the proviso "in comparison to someone of your age; would you say health is excellent, very good, good, fair, poor or don't know?". However, this has not been done in the WVS and during the modelling stage; therefore, it has to condition on age.

The WVS is a world-wide investigation of socio-cultural and political change. It consists of representative national surveys of the basic values and beliefs of the general public in 69 countries. It builds on the European Values Surveys, first carried out in 1981. A second wave of surveys, designed for global use, was completed in 1990-1991, a third wave was carried out in 1995-1996 and a fourth wave took place in 1999-2001. For each country there are interviews with a representative national sample of at least 1,000 people, which are weighted to reflect each country's population.[1] The countries surveyed represent almost 80 percent of the world's population. Further details are given at http://wvs.isr.umich.edu, and http://www.worldvaluessurvey.org, websites and in the major book by Inglehart (1997). The first three waves have been obtained from the UK Data Archive; the 1999-2001 wave was available in April 2004, from a CD ROM accompanying the book *Human Beliefs and Values* (Inglehart *et al.*, 2004).

The key feature of the WVS data set is that the data are available for individuals and has not been aggregated. More importantly, a single survey questionnaire has been used to collect the data from a large number of countries according to scientific (random) sampling procedures. Data was obtained not only on individual self-rated health, but also on household income, social-demographic variables (i.e. age, sex and marital status), and other variables pertinent to the Wilkinson hypothesis, such as happiness and life satisfaction, (these will be employed as the multivariate outcomes in Chapter Five). Consequently, this data provides an ideal empirical test of the relative-income hypothesis.

When studying Wilkinson's income inequality hypothesis, the effect of a range of individual characteristics and income need to be taken into account such as age, gender, and married status. Age and gender structure of a population are key demographic factors

---

[1] These sample weights will be used in the multilevel analysis.

as well as married status, which have been frequently considered in previous health inequality studies (Table 2.1 in Chapter Two). Respondents' current marriage status is categorized into seven conditions (married, live together as married, divorced, separated, widowed, single and unknown). Individual income is measured on a decile scale of income distribution of household counting all wages, salaries, pensions and other incomes before taxes and other deductions. These are all included in the WVS data set; and the Appendix 1 lists the exact questions used in the survey.

4.2.2 Income and income inequality of countries

There are two global data sets on country GDP and income inequality, which are needed to evaluate Wilkinson's income inequality hypothesis. Here I will briefly give a description of the country level income and income inequality indices, and their linkages with the WVS as differentiated from Chapter Three, without overly replicating the same information. The World Bank website provides international country-level income data on GDP per capita (purchase power parity in 2004 US dollars) on an annual basis, from 1970 to 2000 for some 232 countries. This is clearly an ideal source for wave-specific data on a country's absolute income on a comparable basis.

One source of inequality data is the Deininger and Squire (D&S) information provided at the World Bank website, which has often been used as the major source of data in previous studies. In practice, using the data as the main source for all countries found in the WVS would result in a lack of high quality information. Consequently, much of the richness of the individual data of the WVS would have to be omitted, if this was the only source of information that could be used. The second source of income inequality data comes from the UTIP-UNIDO project at the University of Texas. This has some 3,200 observations over 36 years (1963-1999).

All four data sets are being able to be merged so that there is both micro and macro information for 1981, 1991, 1995-7, 1999-2000. In some instances, it is not able to link inequality data to be contemporaneous with the year of the WVS. When this is the case, a lag of up to 5 years is allowed before treating the data as missing. Table 4.1 gives the countries and wave data with sample numbers that form the basis of the cleaned and linked data. Figure 4.1 shows considerable variation between both wave and country when absolute country income (GDP) is plotted against income inequality (EHII, UTIP-UNIDO index). Consequently, this should allow a good test of the Wilkinson hypothesis as overall, countries experience considerable variation in income and/or income inequality.

4.2.3 Preparing to model

In order to facilitate comparisons with previous studies, the dependent variable, self-rated health, was consolidated from five categories into three (Good, Fair and Poor) with Good arbitrarily chosen as the base. In other words, I analysed the underlying probability of reporting 'fair' or 'poor' health in comparison to being in 'good' health. The study design is not a full-panel one (in which the same individuals are repeatedly measured), but a repeated cross-sectional design with new cross-sectional samples of people being taken at each wave. The structure (see Figure 4.2) is therefore three levels: 171,264 individuals at level one, who are nested within four waves at level two nested within 69 countries at level three, and this structure, has to be taken into account during the analysis.

The possible independent variables at the individual level are social-demographic factors (age, sex and marital status) and an individual income variable, while at the wave level measures for relative and absolute income for countries in each wave have been included. Age is the only continuous predictor with a mean of 40 across all countries and waves.

Marital status is consolidated from seven into four categories – couple (combined married couple and live-as-married couple), single, unknown and SWD (separated/widowed/divorce as a group). Individual income is divided into five groups (1&2, 3&4, 5&6, 7&8 and 9&10 of income scale) plus an unknown group.

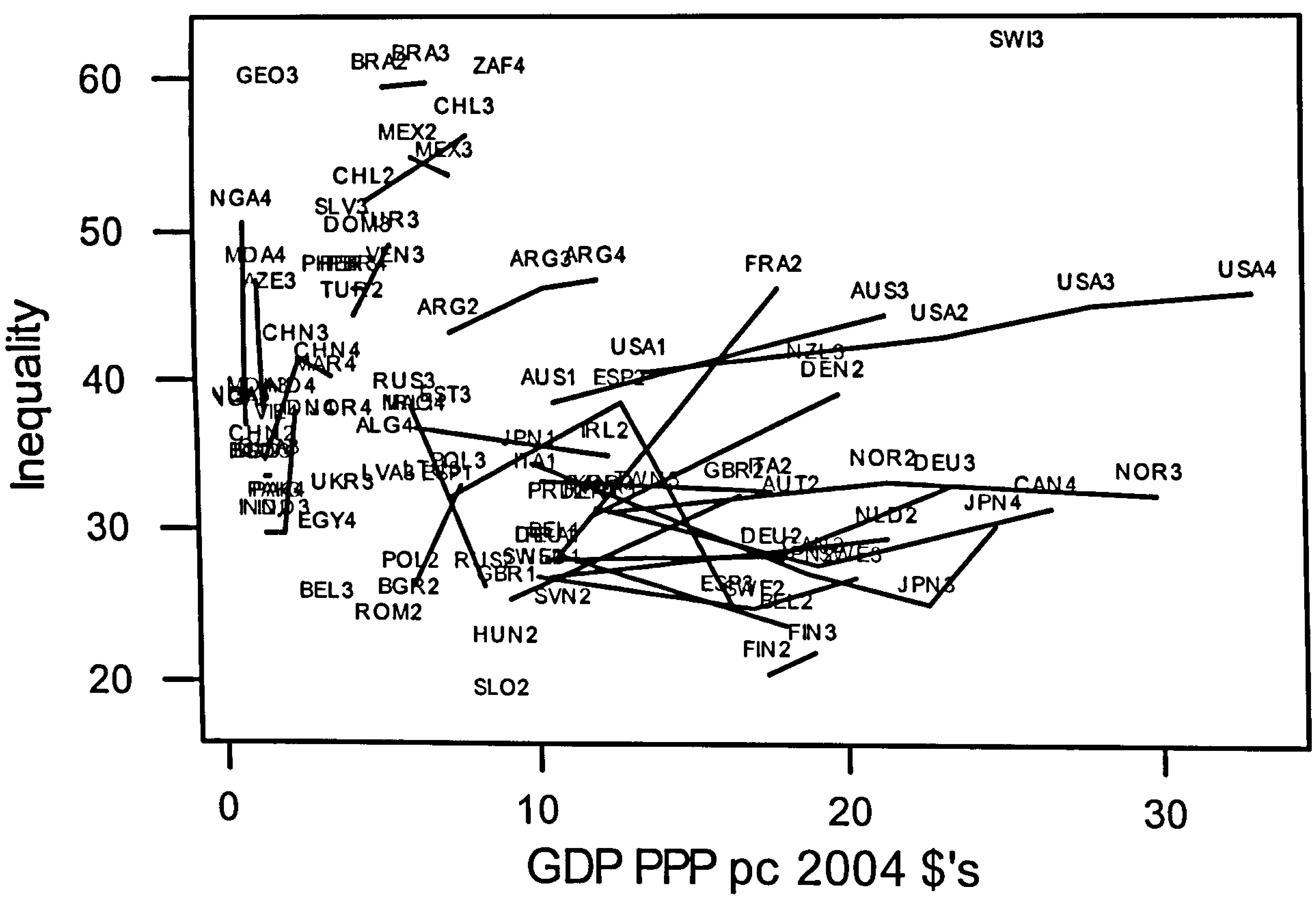
Table 4.1 The structure of the World Values Surveys

Country	Wave	Sample Size	Self-rated Health	Age	Sex	Marital status	Family Income	GDP	UTIP
Albania	W4	1000	*	*	*	*	*	3.45	42.77
Algeria	W4	1282	*	*	*	*	*	5.25	38.09
Argentina	W2	1002	*	*	*	*	*	7.31	43.82
Argentina	W3	1079	*	*	*	*	*	10.37	43.25
Argentina	W4	1280	*	*	*	*	*	12.14	43.86
Armenia	W3	2000	*	*	*	*	*	1.66	54.74
Australia	W1	1228	*	*	*	*	*	10.59	30.64
Australia	W3	2048	*	*	*	*	*	21.33	37.02
....									
Uganda	W4	1002	*	*	*	*	*	1.19	No data
Ukraine	W3	2811	*	*	*	*	*	3.95	38.48
Ukraine	W4	1207		*	*	*	*	3.76	41.85
Tanzania	W4	1171	*	*	*	*	*	0.72	No data
Uruguay	W3	1000	*	*	*	*	*	8.08	44.12
USA	W1	2325	*	*	*	*	*	13.54	35.96
USA	W2	1839	*	*	*	*	*	23.13	36.52
USA	W3	1542	*	*	*	*	*	27.82	36.97
USA	W4	1200	*	*	*	*	*	32.94	43.07
Venezuela	W3	1200	*	*	*	*	*	5.64	46.83
Venezuela	W4	1200		*	*	*	*	5.48	48.76
VietNam	W4	995	*	*	*	*	*	1.88	No data
Zimbabwe	W4	1002	*	*	*	*	*	2.8	45.94

The indication * means this question has been asked in the interview with sufficient answers for different individuals in different countries

There is also considerable 'missingness' in the data, with few countries being measured in all four waves. In practice, it required considerable cleaning of the available data e.g. household income should have been recorded as deciles categories, but for some countries it was actually given as a real currency value.

Figure 4.1 Changing GDP and Inequality: Country and Wave



NB each country is shown by a 3 letters code, followed by a wave index; for example, JPN4 represents Japan at wave 4.

Figure 4.2 Non panel-like but repeated cross-sectional multilevel data structure with individual nested within wave nested country

*W1: 1981, W2: 1990, W3: 1995-1997 and W4: 1999-2001. $J_{1\text{-}10}$ indicates the residences of Japan as well as for U for United Kingdom and T for Taiwan.

4.3 MODELLING FRAMEWORK

4.3.1 Modelling continuous variables with a single-level model

Multilevel modelling will again be the main method of analysis as it allows the differentiation between main and interaction effects for group and individual level variables, and simultaneous analysis at a number of levels. At the outset, the underlying structure is data on three levels: individual self-rated health outcomes at level one; nested within different waves at level two; nested within countries at level three. In order to model self-reported health, a number of predictor variables from both the within-nation level and between-nation level will be included. The required model is a complex one and here a range of models of increasing complexity and faithfulness to the data structure will sequentially be discussed. It will begin by ignoring the multilevel structure and the repeated cross-sectional nature of the WVS design with response of more than two categories. A straightforward micro model at the individual level can then be specified as follows:

$$E(\text{Self-rated health}_i) = \beta_0 X_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} \qquad (4.1)$$

in which the Expectation (E) of self-rated health for individual i is a function of individual's income (X_{1i}), age (X_{2i}) and sex (X_{3i}) of the respondent. So for the moment self-rated health is an underling continuous variable; income and age are also continuous but gender is a binary discrete variable. When one of the predictors in the model is a quantity such as age in which 0 is an absurd value, the intercept (which gives estimated the mean self-rated health) will be un-interpretable. To avoid this, it is a good idea to re-express such quantitative predictors about some convenient value near the middle of their ranges – for our age and income variables, we could use the age relative to 40 years old and mean income; say ten thousands US dollars, respectively.

Moreover, it may be noted that the two genders have been accounted by using a single dummy variable. It takes the value 1 for all the females and the value 0 for all the males. The coefficient; β_3 would then provide the differences between the corresponding group means (females) and the mean of the base or reference group, in this case male (the one for which the gender dummy variable is zero). It should also be noted that the choice of reference group is a matter of convenience – it does not affect the estimates of the differences between the groups. With this specification, the intercept β_0 is the expected average self-rated health for the base category, a 40-year-old male with average income earning of 10k US dollars.

4.3.2 Modelling continuous variables with a multilevel structure

The number of predictors can also readily be extended to take account of other individual factors, but our main focus will remain on individual income, and any other variables will be included as 'controls'. We can specify the model further as follows:

$$E(\text{Self-rated health}_{ij}) = \beta_{0j}X_0 + \beta_{1j}X_{1ij} + \beta_2 X_{2ij} + \beta_3 X_{3ij} \tag{4.2}$$

A key element of this specification is that the intercept is indexed by j for a country, as is the slope term associated with individuals' income. That means there is an effect of living in a country, and the relationship between health and individuals' income is allowed to vary between countries. Therefore it is needed to specify a pair of macro-models, for the random intercepts and slopes, to specify these between-country variations:

$$\beta_{0j} = \beta_0 + (\mu_{0j})$$

$$\beta_{1j} = \beta_1 + (\mu_{1j})$$

The term β_0 represents the average score on self-rated health across all individuals and all countries for the base category, a middle-income 40-year-old men; μ_{0j} is the differential for belonging to country j. Similarly, while β_1 is the general relationship across all countries between self-rated health and individuals' income, the μ_{1j} are the differential slopes. If this is positive, then individual income in that country has a strong impact on health, if it is negative the relation will be weaker and flatter. There are two random terms at the country level and we can summarize their distribution as a joint multivariate Gaussian distribution:

$$(\mu_{0j,} \mu_{1j}) \sim N\begin{bmatrix} \sigma^2_{\mu 0} & \sigma_{\mu 01} \\ \sigma_{\mu 01} & \sigma^2_{\mu 1} \end{bmatrix}$$

The off-diagonal covariance term is important. For example, if $\sigma_{\mu 01}$, the association between the differential intercept and the differential slope for income, is positive, the variance between countries will be greatest for individuals with high individual income (a positive μ_{0j} being associated with a positive μ_{1j}). If none of these variance-covariance terms are significantly different from zero, the country a respondent lives in has no effect on self-rated health once gender, age and individual income has been taken into account. Thus, if none of the higher-level terms are significant all the remaining variation is at the individual level; the country-context has no effect whatsoever.

If this is the case, the Wilkinson's hypothesis will have fallen at the first hurdle. If significant variation is found, however, it does not automatically validate the relative income hypothesis. It then has to assess the relative role of average income and income inequality in accounting for any between-country variation that has been found. Thus revised macro models with these additional variables have to be specified. That is macro country variables will be included to try and account for variation between countries after taken account of individuals' income. These macro variables cannot account for individual variation, as they are constant at that level.

4.3.3 Modelling cross-level interactions

The macro equations can be specified as follows:

$$\beta_{0j} = \beta_0 + \alpha_1 w_{1j} + \alpha_2 w_{2j} + (\mu_{0j})$$

$$\beta_{1j} = \beta_1 + \alpha_3 w_{1j} + \alpha_4 w_{2j} + (\mu_{1j})$$

where w_{1j} is average income for country j; and w_{2j} is the inequality. Combining the micro and polynomial macro equations, results in a two-level overall model:

$$\begin{aligned} E(\text{Self-rated health}_{ij}) = {} & \beta_0 X_{0ij} + \beta_1 X_{1ij} + \beta_2 X_{2ij} + \beta_3 X_{3ij} + \alpha_1 w_{1j} X_{0ij} + \alpha_2 w_{2j} X_{0ij} \\ & + \alpha_3 w_{1j} X_{1ij} + \alpha_4 w_{2j} X_{1ij} + (\mu_{0j} X_{0ij} + \mu_{1j} X_{1ij}) \end{aligned} \tag{4.3}$$

In terms of the relative income hypothesis, the key parameters are β_0, β_1, α_2 and α_4 which define cross-level interaction between country inequality, individual income and self-rated health. In effect it is needed to fit a 'surface' between the variables country inequality, discrete individuals' income groups and self-rated health. Figure 4.3 shows some possible results where there are and there are not substantial effects for each parameter. There are five 3D models of relationships among individual's self-rated health (Z-axis), individual's income (Y-axis) and country's income inequality (EHII, X-axis). These three coordinates' axes are the configuration of the surface.

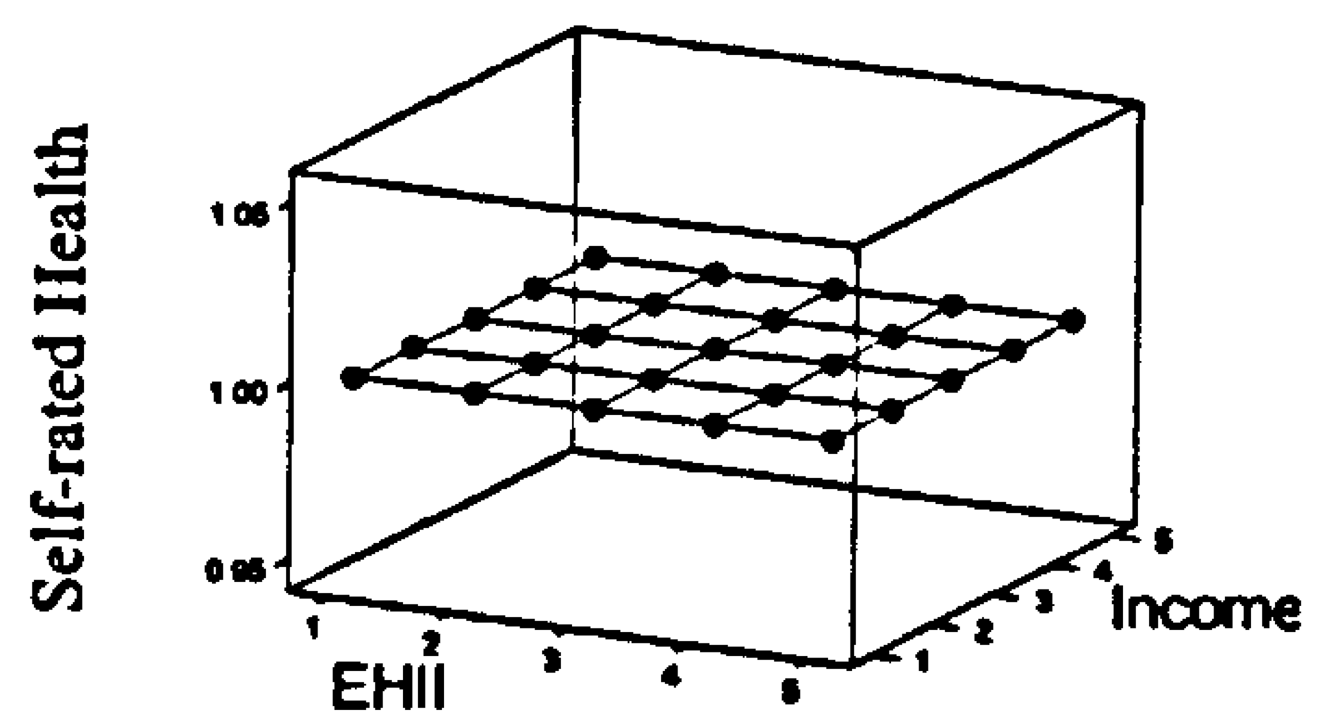
In 4.3a), no matter what value an individual's income or his/her country's income inequality is, the self-rated health is a constant. Redistributing the income distribution of a country or increase an individuals' income will not change their perspective on self-rated health. In Figure 4.3 b), self-rated health improves only when individuals' income increases; no matter how egalitarian a country that they live, rich people have better health than poor everywhere. However, in Figure 4.3c), minimizing a country's income inequality (everyone who lives in that country has the same health condition, no matter how rich you are) will improve their self-rated health.

Moreover, in Figure 4.3 d), on the contrary to Figure 4.3 a), not only individual's income has positive effect but also its degree of income inequality has a negative impact on self-rated health on the condition of no cross-level interaction between them. In other words, people who earn the highest income as well as live in the most egalitarian country benefit from the best self-rated health. Moreover, a wealthy person who lives in a very unequal country will experience better health than a poor person in the most egalitarian country. In the final Figure 4.3 e) individual's income and their country's income inequality have a significant cross-level effect. This means living in different income inequality countries there is different degree impact from individual's income on their self-rated health.

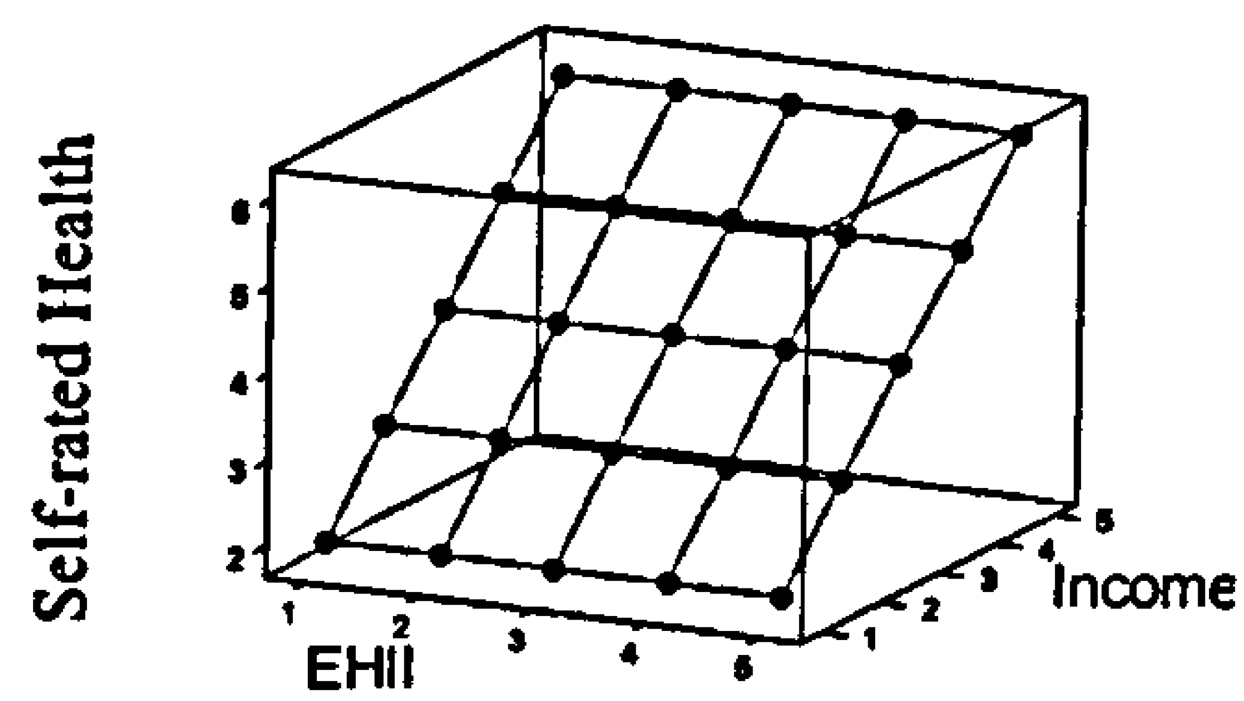
4.3.4 Modelling predictors with a set of discrete categories

In fact, income is not a continuous variable, but is measured as a discrete variable with six categories '1&2' (poorest), '3&4', '5&6', '7&8', '9&10' (wealthiest) and 'unknown-income' in this study. This is modelled by including a set of dummy variable indicators, X_{1ij}, X_{4ij} to X_{7ij}, representing 4 dummy variables contrasted against the base. The middle-income group '5&6' has been chosen as the base. The first of these additional predictors X_{1ij} would take the value 1 for all the people in the income group 1&2 and the value 0 for all the others; the second X_{4ij} the value 1 for all the people in income group 3&4 and the value 0 for all the others, and so on.

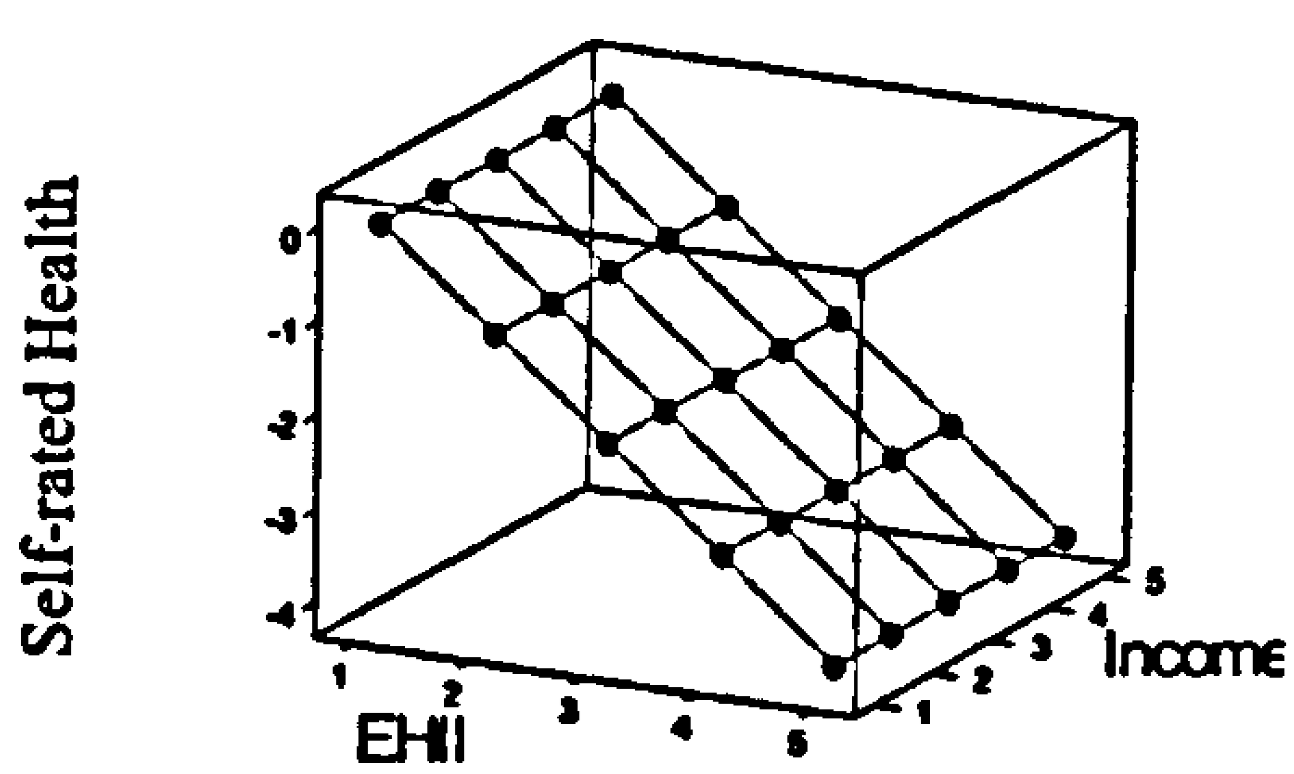
Figure 4.3 Some possible results where there are and there are not a substantial effect for each parameter.



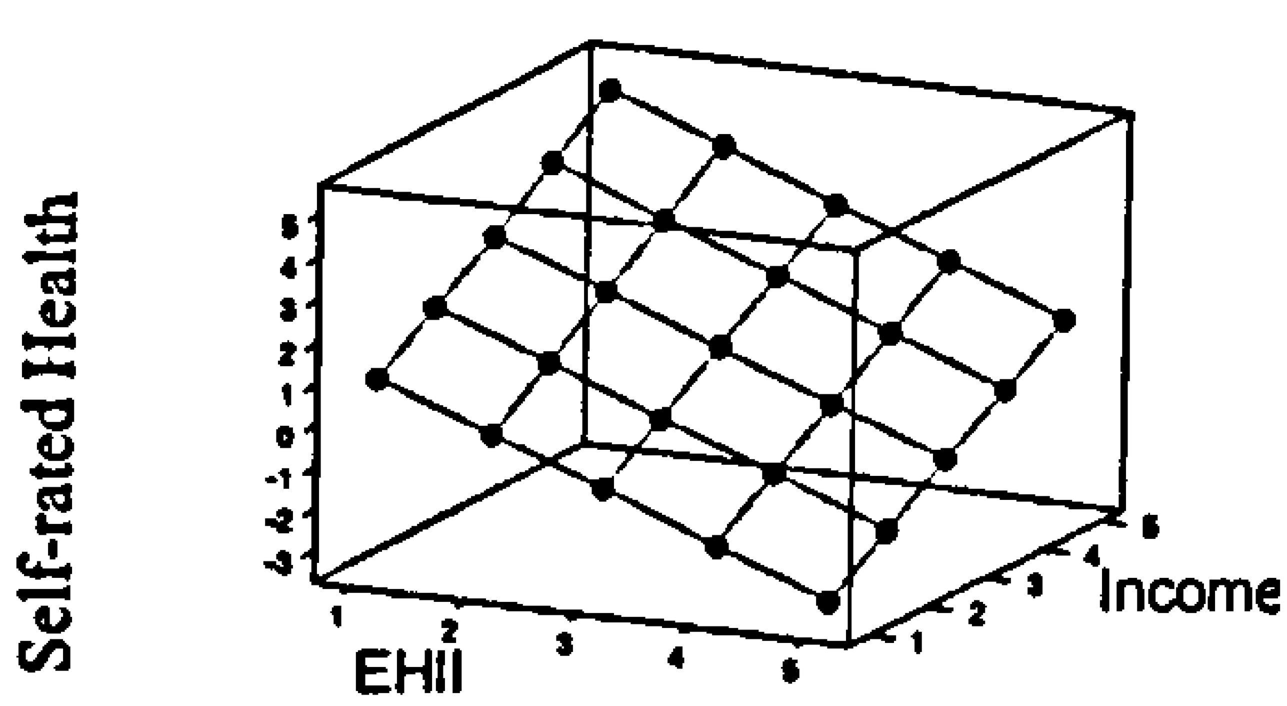
a) No effect of either



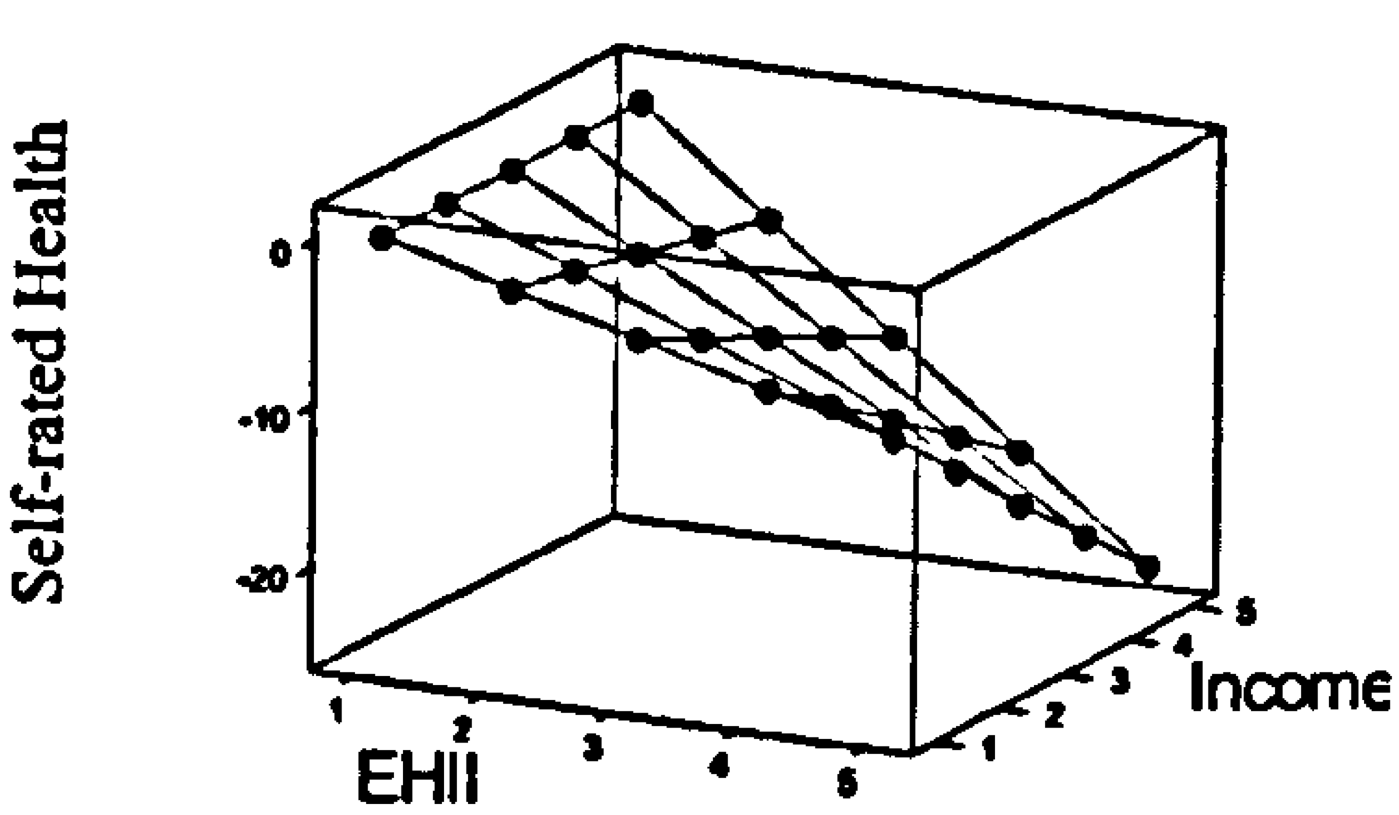
b) Income effect only



c) EHII effect only



d) Both income and EHII effect without cross-level effect



e) Both income and EHII effect with cross-level effect

Turning to the random part specification it would be possible to write macro equations for each income group, but given that there are only 69 countries there would be few degrees of freedom to estimate all the variance covariance involved. Therefore an alternative three-category variable has been created – 'Low-income' (including income groups of 1&2 and 3&4), 'High-income' (which combined income groups of 5&6, 7&8 and 9&10) and 'Unknown-income' (income group unknown) and included this reduced specification to the random part of the model. 'High-income' is the base here. Therefore the specification of the combined micro-macro model is:

$$\begin{aligned} E(\text{Self-rated health}_{ij}) = {} & \beta_{0j}X_{0ij} + \beta_1 X_{1ij} + \beta_2 X_{2ij} + \beta_3 X_{3ij} + \beta_4 X_{4ij} + \beta_5 X_{5ij} + \beta_6 X_{6ij} + \beta_7 X_{7ij} \\ & + \alpha_1 w_{1j} X_{0ij} + \alpha_2 w_{2j} X_{0ij} + \alpha_3 w_{1j} X_{1ij} + \alpha_4 w_{2j} X_{1ij} + \alpha_5 w_{1j} X_{4ij} + \\ & \alpha_6 w_{2j} X_{4ij} + \alpha_7 w_{1j} X_{5ij} + \alpha_8 w_{2j} X_{5ij} + \alpha_9 w_{1j} X_{6ij} + \alpha_{10} w_{2j} X_{6ij} \\ & + \alpha_{11} w_{1j} X_{7ij} + \alpha_{12} w_{2j} X_{7ij} \quad (\mu_{0j} X_{0ij} + \mu_{8j} \textbf{Low-income}_{ij} + \\ & \mu_{9j} \textbf{Unknown-income}_{ij}) \end{aligned} \qquad (4.4)$$

where w_{1j} is average income for country j, and w_{2j} is the inequality. β_1 and β_4 are the income differentials for groups 1&2 and 3&4, respectively in comparison to the base; while μ_{0j} is the country differential for high-income group and μ_{8j} is the country differential for low-income category in comparison to the high-income group. To avoid multi-collinearity with the variable representing individual's income, the new dummies for Low- and Unknown- income are only included in the random part of the model. The same principles and interpretations to the continuous individual's income variable will be applied here too. If μ_{8j} and μ_{9j} are positive, then individuals with low-income and unknown income in that country have a strong impact on health respectively, if it is negative the relationship will be weaker and flatter. In terms of the relative income hypothesis, the key parameters are $\beta_0, \beta_1, \beta_{4\sim7}$, α_2, $\alpha_4, \alpha_6, \alpha_8, \alpha_{10}$ and α_{12}

which define cross-level interaction between country inequality, every single category of discrete individuals' income groups and self-rated health.

There are now three random terms at the country level and it is important to understand how these are interpreted as they give us key information about how income effects differ in different countries. We can summarize the distribution of the three random terms as a joint multivariate Gaussian distribution:

$$(\mu_{0j}, \mu_{8j} \text{ and } \mu_{9j}) \sim N\begin{bmatrix} \sigma^2_{\mu 0} & \sigma_{\mu 08} & \sigma_{\mu 09} \\ \sigma_{\mu 08} & \sigma^2_{\mu 8} & \sigma_{\mu 89} \\ \sigma_{\mu 09} & \sigma_{\mu 89} & \sigma^2_{\mu 9} \end{bmatrix}$$

The off-diagonal covariance term is important and plays the same role as continuous individual's income in previous equation. When the predictors are categorical, the variability for a contrasted category depends on the variance of based category, plus variance of contrast category, plus twice covariance of the base and contrast. For example, the between-country variance of individuals with low income is "$\sigma^2_{\mu 0} + \sigma^2_{\mu 8} + 2^* \sigma_{\mu 08}$", the variance of 'high-income' plus the variance of 'low-income' and plus twice the association between the differential intercept and the differential slope for 'low-income', and vice versa for other groups. Again if none of these variance-covariance terms are significantly different from zero, the country a respondent lives has no effect on self-rated health once gender, age and individual income has been taken into account. Thus, if none of the higher-level terms are significant, all the remaining variation is at the individual level; the country-context has no effect whatsoever.

4.3.5 Modelling time as a level and as a fixed part difference

The next extension will be to put time in the model as a level, since the WVS has a repeated cross-sectional structure in which a different non-panel sample has been drawn on each occasion that the survey was undertaken. This will allow an assessment of trend of the health-inequality relation over time taking account of individual income. The discrete time variable is also included in the fixed part as a three dummies variable with wave 1 as the base as well as a level, which have individuals nested within this level, and this level nested within countries. The three-level model can be defined as follows where i is individual, j is wave and k is country:

$$\begin{aligned}E(\text{Self-rated health}_{ijk}) = {} & \beta_{0jk}X_{0ijk} + (\beta_{10}X_{10ijk} + \beta_{11}X_{11ijk} + \beta_{12}X_{12ijk}) + \beta_1X_{1ijk} + \beta_2X_{2ijk} + \\ & \beta_3X_{3ijk} + \beta_4X_{4ijk} + \beta_5X_{5ijk} + \beta_6X_{6ijk} + \beta_7X_{7ijk} + \alpha_1w_{1k}X_{0ijk} + \\ & \alpha_2w_{2k}X_{0ijk} + \alpha_3w_{1k}X_{1ijk} + \alpha_4w_{2k}X_{1ijk} + \alpha_5w_{1k}X_{4ijk} + \alpha_6w_{2k}X_{4ijk} \\ & + \alpha_7w_{1k}X_{5ijk} + \alpha_8w_{2k}X_{5ijk} + \alpha_9w_{1k}X_{6ijk} + \alpha_{10}w_{2k}X_{6ijk} \\ & + \alpha_{11}w_{1k}X_{7ijk} + \alpha_{12}w_{2k}X_{7ijk} + (v_{0jk}X_{0ijk}) + (\mu_{0k}X_{0ijk} \\ & + \mu_{8k}\text{Low-income}_{ijk} + \mu_{9k}\text{Unknown-income}_{ijk}) \qquad (4.5)\end{aligned}$$

The new parameters in the fixed part of the model are associated with the dummy variables X_{10ijk}, X_{11ijk} and X_{12ijk}, representing waves two, three and four respectively; These parameters give the differentials between each subsequent wave and the first base wave of 1981. However, they are not the same respondents in each time wave nor are they always the same countries, so that one needs to careful with the interpretation of health trend through time. The above equation also now contains the subscript ijk in the random part, representing individuals within waves within countries.

4.3.6 Modelling a discrete outcome variable

Finally it comes to the last extension that required. So far the responses have been treated as if it was a continuous variable, whereas in fact it is a set of categories: good, fair, poor health. One approach would be to fit a linear probability model, but there are three problems with ignoring the discrete nature of the dependent variable:

1 Nonsensical values where fitted proportion are unbounded and can get predictions outside the range 0 to 1;

2 We can anticipate a non-linear relationship between response and continuous predictors such as age as the limits of the underlying probability of 0 to 1 are approached;

3 Inbuilt heterogeneity such that there is less variation as approach bounds of 0 and 1, that is variation between individuals is not Gaussian but multinomial and depends on the underlying predicted probability.

These problems can be resolved as follows:

For problems 1 and 2, a non-linear logit transformation of the response in the micro-model is used. However, it can readily be transformed back to proportions or probabilities for interpretation. For problem 3, a 'weighted' estimation is used to create 'level-1 weights' for each response whereby each individual will have probabilities of choosing each category that sum to 1; the weight is chosen so that if there is exact multinomial residual distribution the associated variance will be 1.[2] This multinomial model can be extended to multilevel multinomial models (based on a logit-link function), with individuals in a particular higher-level unit sharing the

[2] These weights are not to be confused with sampling weights which are also used in the models to make the data nationally representative.

same underlying category probabilities. Therefore, the real level one (individuals) becomes level two due to treating the responses nested under individuals as 'level-one weights'. For the jth individual at level two which been interviewed in the kth wave at level three nested within the lth country at level four (see the structure in Figure 4.4), the probability the individuals self-rated their health as one of the terms (poor, fair or good) at level one is π^{h}_{ijkl}, h = poor, fair and good, where $\sum_h \pi^{(h)}_{ijkl} = 1$. Treating self-rated good health as the base group, with only three categories of responses in total, we require two models for the log ratios $\log\left(\frac{\pi^{poor}_{ijkl}}{\pi^{good}_{ijkl}}\right)$ and $\log\left(\frac{\pi^{fair}_{ijkl}}{\pi^{good}_{ijkl}}\right)$, respectively.

Taking into account the clustering of responses nested within individuals nested within wave nested within country, we have the four-level multinomial logistic regression model with two "wings"; treating each as a separate category allows for different effects in each wing of the model for poor/good compared to fair/good as follows:

$$\begin{aligned}
E(\mathrm{Log}\left(\frac{\pi^{poor}_{ijkl}}{\pi^{good}_{ijkl}}\right)) = {} & \beta_{0kl}X_{0ijkl} + (\beta_{10}X_{10ijkl} + \beta_{11}X_{11ijkl} + \beta_{12}X_{12ijkl}) + \beta_1X_{1ijkl} + \beta_2X_{2ijkl} + \\
& \beta_3X_{3ijkl} + \beta_4X_{4ijkl} + \beta_5X_{5ijkl} + \beta_6X_{6ijkl} + \beta_7X_{7ijkl} + \alpha_1w_{1l}X_{0ijkl} + \\
& \alpha_2w_{2l}X_{0ijkl} + \alpha_3w_{1l}X_{1ijkl} + \alpha_4w_{2l}X_{1ijkl} + \alpha_5w_{1l}X_{4ijkl} + \alpha_6w_{2l}X_{4ijkl} \\
& + \alpha_7w_{1l}X_{5ijkl} + \alpha_8w_{2l}X_{5ijkl} + \alpha_9w_{1l}X_{6ijkl} + \alpha_{10}w_{2l}X_{6ijkl} + \alpha_{11}w_{1l}X_{7ijkl} + \\
& \alpha_{12}w_{2l}X_{7ijkl} + (v_{0jkl}X_{0ijk}) + (\mu_{0kl}X_{0ijkl} + \mu_{8l}\mathrm{Low\text{-}income}_{ijkl} + \\
& \mu_{9l}\mathrm{Unknown\text{-}income}_{ijkl}) \qquad (4.6.1)
\end{aligned}$$

$$E(\text{Log}\left(\frac{\pi_{ijkl}^{poor}}{\pi_{ijkl}^{good}}\right)) = \beta_{0kl}X_{0ijkl} + (\beta_{10}X_{10ijkl} + \beta_{11}X_{11ijkl} + \beta_{12}X_{12ijkl}) + \beta_1X_{1ijkl} + \beta_2X_{2ijkl} + \beta_3X_{3ijkl} + \beta_4X_{4ijkl} + \beta_5X_{5ijkl} + \beta_6X_{6ijkl} + \beta_7X_{7ijkl} + \alpha_1w_{11}X_{0ijkl} + \alpha_2w_{21}X_{0ijkl} + \alpha_3w_{11}X_{1ijkl} + \alpha_4w_{21}X_{1ijkl} + \alpha_5w_{11}X_{4ijkl} + \alpha_6w_{21}X_{4ijkl} + \alpha_7w_{11}X_{5ijkl} + \alpha_8w_{21}X_{5ijkl} + \alpha_9w_{11}X_{6ijkl} + \alpha_{10}w_{21}X_{6ijkl} + \alpha_{11}w_{11}X_{7ijkl} + \alpha_{12}w_{21}X_{7ijkl} + (v_{0jkl}X_{0ijk}) + (\mu_{0kl}X_{0ijkl} + \mu_{8l}\text{Low-income}_{ijkl} + \mu_{9l}\text{Unknown-income}_{ijkl}) \quad (4.6.2)$$

In the random part, there is now a more complex covariance structure with the two "wings" of the higher level, each having their own variance and being linked by a co-variance as specified in Table 4.2.

These two variance and covariance matrices provide the information on the variability of the effects of any of the variables across countries or between waves. The six terms on the diagonal are the random effects of the two 'intercepts' associated with the base groups and the random effects of the four 'slopes' associated with the variable 'low-income' and 'unknown-income', respectively. The co-variances are the associations between proportion of poor/good and fair/good of intercept and slopes, between waves. For example, $\sigma_{\mu 0}^{(poor,fair)}$ shows the covariance between the log-odds of reporting poor and fair health. At the country level, a positive estimate implies that a country with a high proportion of reporting poor health tends to have a high proportion of reporting fair health.

Figure 4.4 Non panel-like but repeated cross-sectional multilevel data structure with individuals nested within waves nested countries

*W1: 1981, W2: 1990, W3:1995-1997 and W4:1999-200. $J_{1\text{-}10}$ indicate the residences of Japan as well as for U for United Kingdom and T for Taiwan. P, F and G means each individual self-rated she or herself health as poor, fair and good, respectively.

Table 4.2 Variance matrix at the country level and wave level

At the country level

	μ_{0l}^{poor}	μ_{0l}^{fair}	$\mu_{low-incomel}^{poor}$	$\mu_{low-incomel}^{fair}$	$\mu_{unknown-incomel}^{poor}$	$\mu_{unknown-incomel}^{fair}$
μ_{0l}^{poor}	$\sigma_{\mu 0}^{2(poor)}$					
μ_{0l}^{fair}	$\sigma_{\mu 0}^{(poor,fair)}$	$\sigma_{\mu 0}^{2(fair)}$				
$\mu_{low-incomel}^{poor}$	$\sigma_{\mu 0 low-income}^{(poor,poor)}$	$\sigma_{\mu 0 low-income}^{(fair,poor)}$	$\sigma_{\mu low-income}^{2(poor)}$			
$\mu_{low-incomel}^{fair}$	$\sigma_{\mu 0 low-income}^{(poor,fair)}$	$\sigma_{\mu 0 low-income}^{(fair,fair)}$	$\sigma_{\mu low-income}^{(fair,poor)}$	$\sigma_{\mu low-income}^{2(fair)}$		
$\mu_{unknown-incomel}^{poor}$	$\sigma_{\mu 0 unknown-income}^{(poor,poor)}$	$\sigma_{\mu 0 unknown-income}^{(poor,fair)}$	$\sigma_{\mu unknown-income, low-income}^{(poor,poor)}$	$\sigma_{\mu unknown-income, low-income}^{(poor,fair)}$	$\sigma_{\mu unknown-income}^{2(poor)}$	
$\mu_{unknown-incomel}^{fair}$	$\sigma_{\mu 0 unknown-income}^{(poor,fair)}$	$\sigma_{\mu 0 unknown-income}^{(fair,fair)}$	$\sigma_{\mu unknown-income, low-income}^{(fair,poor)}$	$\sigma_{\mu unknown-income, low-income}^{(fair,fair)}$	$\sigma_{\mu unknown-income}^{(poor,fair)}$	$\sigma_{\mu unknown-income}^{2(fair)}$

At the wave level:

	v_{0k}^{poor}	v_{0k}^{fair}
v_{0k}^{poor}	$\sigma_{v0}^{2(poor)}$	
v_{0k}^{fair}	$\sigma_{v0}^{(poor,fair)}$	$\sigma_{v0}^{2(fair)}$

4.3.7 Estimation procedures

There are a number of estimation strategies for estimating multilevel models with discrete response Goldstein (2003) proposed fitting multiple categorical responses models with multinomial error distributions using IGLS (iterative generalized least squares). However there are two major drawbacks with this approach. First, it is known that maximum likelihoods estimators like IGLS underestimate the size of the random effects (Rodriguez and Goldman, 2001). Moreover, there is no convenient overall summary measure of goodness of fit that allows model comparison. Another and improved estimation strategy for discrete and complex models is the MCMC (short for Markov Chain Monte Carlo) methodology.

Instead of finding simple point estimates for the parameters of interest, the MCMC methodology makes a large number of simulated random draws from the joint posterior distribution of all the parameters. It then uses these random draws to form a summary of the underlying distribution (The book by Gilks et al., 1996 gives more in-depth material.). The models in this chapter have been fitted in the multilevel; modeling software MLwiN (Rasbash *et al.,* 2004). MLwiN uses IGLS to derive starting parameters and then a combination of two MCMC procedures, Gibbs sampling and Metropolis-Hastings sampling. In Bayesian statistics, every unknown parameter must have a prior distribution. With all the fitted models, the prior distributions were set to be 'flat' or 'diffuse' for all the parameters implying, that we do not have a strong preference for any particular values of the estimates before fitting the model.

It is a general practice to throw away the first n values generated to allow the Markov chain to approach its equilibrium distribution, namely the joint posterior distribution of interest. These n values are known as a 'burn-in' and I have used the default of 500 such simulations. Then continue generating values after the 'burn in' for other m simulations; these m values are then averaged to give estimates of the parameter of interest. Posterior standard deviations (like frequentist standard errors) for the estimates can be obtained by calculating the standard deviations of the m values.

For all the model results given below, the IGLS (iterative generalized least squares) algorithm of Goldstein (1986) is used to get an initial estimate of the parameters, then a burn-in of 500 and then 50,000 further estimates to explore the parameter space. I also inspected the 'traces' of key parameters (a plot of the parameter against simulation number) to ensure that convergence had been achieved. All models were estimated using the logit link (logarithm of the odds) function. The exponentiated coefficients from a logit model are interpreted as odds ratios (OR). Results will be reported as 'significant' if the estimates are more than twice their estimated empirical standard error.

4.3.8 Model Comparison

The Deviation Information Criterion (DIC) which was introduced by Spiegelhalter and his colleague (2002) has been used to evaluate how well the model has fitted the data in comparison with another model. The DIC is an extension of the Akaike's Information Criterion (AIC) the latter being an estimate of the expected relative distance between the fitted model and the unknown true mechanism that generated the observed data. The DIC diagnostic can only be calculated from the chains produced by an MCMC run.

The DIC assesses the badness of a fit to the model, taking into account the complexity of the model. The diagnostic DIC = Deviance + 2*pD where pD is the 'effective number of parameters', which can be calculated from the chain as the difference between the mean deviance ($\overline{D}$) in the chain and the deviance at the mean values for the parameters (D($\overline{\theta}$)). In random effect models the 'effective number of parameters' is less than an equivalent fixed effect model would have, due to the additional distributional assumption for the random effects. That is while we might anticipate 5 extra degrees of freedom being required for 5 contrasted income categories, somewhat less than 69 degrees of freedom would be consumed in fitting the between country differentials as they are assumed to have come from an overall distribution.

For Multinomial models the deviance formula is $D = -2\sum_{i}\sum_{j=1}^{t}\left[I(y_i = j)\hat{\pi}_i^{(j)}\right]$ where I() is an indicator function which returns 1 if the condition is satisfied, i.e. if individual i is in self-rated health group j. $\hat{\pi}_i^{(j)}$ is the estimated probability of being in category j for individual i. This means that the MCMC engine in MLwiN calculates the estimated probabilities as part of the DIC diagnostic command (Browne, 2004, p266-267). The DIC diagnostic will then give a single number for each model with the smallest value representing the best model, having taken account of model estimated complexity.

4.3.9 Analysis strategy

In the previous section a range of models that could be fitted to the data have been considered, and the organization was primarily to explain the underlying character of the model in an approachable way. In this section, a serious of models will be built in terms of increasing complexity to test various aspects of the Wilkinson's hypothesis.

Model 1: this is the so called null model (empty, without including any predictors in the fixed part of the model) with the exception of a set of dummy variable indicating wave (wave 1 as reference) in the fixed part of the model. In the random part, the intercept (associated with the constant term) is allowed to be vary both at levels four (country) and three (between cohorts within countries). This model provides an estimate of the global pattern of self-rated health at baseline (1981), the differential from baseline at each subsequent wave, and the between-country and within-country-between-wave variations in self-rated health. The relative size of these higher-level variances can be compared in subsequent models as predictor variables are introduced. The models are specified as follows:

$$E(\log(\pi_{ijkl}^{poor} / \pi_{ijkl}^{good})) = \beta_{0kl}X_{0ijkl} + \beta_1 W2_{ijkl} + \beta_2 W3_{ijkl} + \beta_3 W4_{ijkl} + (\mu_{0l}X_{0ijkl} + \nu_{0kl}X_{0ijkl})$$

$$E(\log(\pi_{ijkl}^{fair} / \pi_{ijkl}^{good})) = \beta_{0kl}X_{0ijkl} + \beta_1 W2_{ijkl} + \beta_2 W3_{ijkl} + \beta_3 W4_{ijkl} + (\mu_{0l}X_{0ijkl} + \nu_{0kl}X_{0ijkl})$$

where $\log(\pi_{ijkl}^{poor} / \pi_{ijkl}^{good})$ is the probability of reporting poor health in logit and indicator i is used here to differentiate it from $\log(\pi_{ijkl}^{fair} / \pi_{ijkl}^{good})$ for reporting fair health of individual j nested within time point k nested within country l. For interpretation, odds ratio (OR) and the probability of reporting fair and poor health in certain conditions are employed, respectively. The constant is represented by X_{0ijkl}; $W2_{kl}$, $W3_{kl}$ and $W4_{kl}$ are the dummy time point variables (W1 as the base) of each survey wave. The fixed part parameters, which show the overall relationship between both individual and country predictors to our response outcomes now have the following meaning: β_{0kl}, when transformed, represents the global mean probability of poor health of the reference group which in this model is everyone in the survey in 1981; $\beta_1, \beta_2, \beta_3$ define either the increasing or decreasing on log-odds of reporting poor health in each survey time point after 1981. In the random part, each parameter is the variation between countries and waves in self-rated health that cannot be accounted for by these included factors. So μ_{0jk} is the country differential in 1981 and v_{0k} is the difference between waves within each country.

Model 2: builds on model 1 by including all the individual predictors in the fixed part of the model. It models how much an individual's income affects their perspective that their own health was poor after having controlled for all other individual characteristics – age, sex and married status. Age is a continuous variable and has been centred on its mean 40. Other base categories are 'couple' for the married status variable and '5&6 decile of income scale' for income variable. In the random part, the effect for the individual income variable is allowed to be varied at level four. This is done as discussed earlier for the combined groups: low income, high income and unknown income. Consequently, in moving from model 1 to model 2, the contextual variation between countries in terms of self-rated health was estimated before and after taking into account the compositional effect of individual demographic and income variables. The models are specified as follows, with the new terms shown in bold:

$$E(\log(\pi_{ijkl}^{poor} / \pi_{ijkl}^{good})) = \beta_{0kl}X_{0ijkl} + \beta_1 W2_{ijkl} + \beta_2 W3_{ijkl} + \beta_3 W4_{ijkl} + \boldsymbol{\beta_4(Age\text{-}40)_{ijkl}} + \boldsymbol{\beta_5 Female_{ijkl}} + \boldsymbol{\beta_6(Age\text{-}40)*Female_{ijkl}} + \boldsymbol{\beta_7 Msingle_{ijkl}} + \boldsymbol{\beta_8 Mseperated/window/divoced_{ijkl}} + \boldsymbol{\beta_9 MUnknown_{ijkl}} + \boldsymbol{\beta_{10} I1\&2_{ijkl}} + \boldsymbol{\beta_{11} I3\&4_{ijkl}} + \boldsymbol{\beta_{12} I7\&8_{ijkl}} + \boldsymbol{\beta_{13} IUnknown_{ijkl}} + (\mu_{0l}X_{0ijkl} + v_{0kl}X_{0ijkl} + \boldsymbol{\mu_{14l} Low\text{-}income_{ijkl}} +$$

μ_{15l}**Unknown-income$_{ijkl}$)**

$$E(\log(\pi^{fair}_{ijkl} / \pi^{good}_{ijkl})) = \beta_{0kl}X_{0ijkl} + \beta_1 W2_{ijkl} + \beta_2 W3_{ijkl} + \beta_3 W4_{ijkl} + \beta_4(\text{Age-40})_{ijkl} + \beta_5\text{Female}_{ijkl} + \beta_6(\text{Age-40})*\text{Female}_{ijkl} + \beta_7\text{Msingle}_{ijkl} + \beta_8\text{Mseperated/window/divoced}_{ijkl} + \beta_9\text{MUnknown}_{ijkl} + \beta_{10}\text{I1\&2}_{ijkl} + \beta_{11}\text{I3\&4}_{ijkl} + \beta_{12}\text{I7\&8}_{ijkl} + \beta_{13}\text{IUnknown}_{ijkl} + (\mu_{0l}X_{0ijkl} + \nu_{0kl} X_{0ijkl} + \mu_{14l}\text{Low-income}_{ijkl} + \mu_{15l}\text{Unknown-income}_{ijkl})$$

where Age_{ijkl} is the age for each individual centred on age 40; β_4 defines the mean log-odds of self-rated poor health for being one year older. F_{ijkl} is an indicator variable identifying females with a 1, males with a 0; β_5 is the gender gap of mean log-odds of self-rated poor health; the difference between males and females. $F*Age_{ijkl}$ is the interaction between the female dummy and continuous age; β_6 is therefore the age * gender interaction. The contrasted categories of marital status are $Mswd_{ijkl}$, $Msig_{ijkl}$ and $Munk_{ijkl}$ (Couple as the base) representing widowed/separate/divorced, single and unknown respond, respectively. Therefore, β_7, β_8, β_9 defines the difference for marital status in contrast to couple (the base). The terms; $I1\&2_{ijk}$, $I3\&4_{ijk}$, $I7\&8_{ijk}$, $I9\&10_{ijk}$ and $Iunk_{ijk}$ (I5&6 as the base) are the income dummies in quartile; β_{10}, β_{11}, β_{12}, β_{13} are the differentials of each income group in contrast to group 5 and 6. So the reference group β_{0kl} is now referring to the middle class 40-year-old married men in the 1981 cohort. In the random part, there are two extra terms from model 1 and 2 that μ_{14l} is the slope differential for low-income and μ_{15l} is the slope differential of unknown-income at the country level, contrasted against the high income group country differential μ_{0l}.

Model 3: builds on model 2, by adding the fixed effect of a country's income on individuals' self-rated poor health. It tests the absolute income hypothesis whether the level of a country's wealthy affects people's health after controlling for their individual socio-demographical and economic characteristics. It also considers whether the income effect differs for countries with GDP above five thousand US dollars (Wilkinson's threshold) from below it. By estimating the relationship between individual's self-rated health and a country's GDP PPP pc 2004$, the model is defined as follows:

$E(\log(\pi_{ijkl}^{poor}/\pi_{ijkl}^{good})) = \beta_{0kl}X_{0ijkl} + \beta_1 W2_{ijkl} + \beta_2 W3_{ijkl} + \beta_3 W4_{ijkl} + \beta_4(\text{Age-40})_{ijkl} + \beta_5 \text{Female}_{ijkl} + \beta_6(\text{Age-40})*\text{Female}_{ijkl} + \beta_7 \text{Msingle}_{ijkl} + \beta_8 \text{Mseperated/window/divoced}_{ijkl} + \beta_9 \text{MUnknown}_{ijkl} + \beta_{10} \text{I1\&2}_{ijkl} + \beta_{11} \text{I3\&4}_{ijkl} + \beta_{12} \text{I7\&8}_{ijkl} + \beta_{13} \text{IUnknown}_{ijkl} + \boldsymbol{\beta_{16}(\text{GDP-10})_{kl}} + \boldsymbol{\beta_{17}(\text{Wilkinson threshold})_{kl}} + \boldsymbol{\beta_{18}(\text{Wilkinson threshold})_{kl}*(\text{GDP-10})_{kl}} + (\mu_{0l}X_{0ijkl} + \nu_{0kl}X_{0ijkl} + \mu_{14l}\text{Low-income}_{ijkl} + \mu_{15l}\text{Unknown-income}_{ijkl})$

$E(\log(\pi_{ijkl}^{fair}/\pi_{ijkl}^{good})) = \beta_{0kl}X_{0ijkl} + \beta_1 W2_{ijkl} + \beta_2 W3_{ijkl} + \beta_3 W4_{ijkl} + \beta_4(\text{Age-40})_{ijkl} + \beta_5 \text{Female}_{ijkl} + \beta_6(\text{Age-40})*\text{Female}_{ijkl} + \beta_7 \text{Msingle}_{ijkl} + \beta_8 \text{Mseperated/window/divoced}_{ijkl} + \beta_9 \text{MUnknown}_{ijkl} + \beta_{10} \text{I1\&2}_{ijkl} + \beta_{11} \text{I3\&4}_{ijkl} + \beta_{12} \text{I7\&8}_{ijkl} + \beta_{13} \text{IUnknown}_{ijkl} + \boldsymbol{\beta_{16}(\text{GDP-10})_{kl}} + \boldsymbol{\beta_{17}(\text{Wilkinson threshold})_{kl}} + \boldsymbol{\beta_{18}(\text{Wilkinson threshold})_{kl}*(\text{GDP-10})_{kl}} + (\mu_{0l}X_{0ijkl} + \nu_{0kl}X_{0ijkl} + \mu_{14l}\text{Low-income}_{ijkl} + \mu_{15l}\text{Unknown-income}_{ijkl})$

Where GDP-10$_{kl}$ refers to the average country income in term of gross domestic product in purchase power parity of 2004 US dollars per capita which has been centred around the global mean of ten thousand dollars. The term β_{16} estimates the effect of increasing by one thousand dollar a country's average income on people's perception of their own health. Wilkinson's threshold is a categorical variable which divides countries into two types; one is the countries with mean income above five thousand dollars in 1990 and the other type is the countries having less than five thousand dollars. The countries below than Wilkinson's threshold have been treated as the base here. The term β_{17} assesses the differential effect of a country having average income of above five thousand dollars in 1990 from that below the threshold.

An interaction term -β_{18}- between GDP and Wilkinson's threshold, allows us to estimate the effect of changing GDP for these two different types of countries affecting people's perception of reporting their health condition. The reference group here is the middle income, 40-year-old married men who live in a country with average income of ten thousand US dollars per person and their countries in 1990 belong to the below five thousands US dollars group in the cohort of 1981.

Model 4 builds on model 3, by considering the fixed effect of the contextual variable, country income inequality on individual's self-rated health and the extent to which it explains the country-level differences. The model can be specified as follows:

$$E(\log(\pi_{ijkl}^{poor} / \pi_{ijkl}^{good})) = \beta_{0kl}X_{0ijkl} + \beta_1 W2_{ijkl} + \beta_2 W3_{ijkl} + \beta_3 W4_{ijkl} + \beta_4(\text{Age-40})_{ijkl} + \beta_5\text{Female}_{ijkl} + \beta_6(\text{Age-40})*\text{Female}_{ijkl} + \beta_7\text{Msingle}_{ijkl} + \beta_8\text{Mseperated/window/divoced}_{ijkl} + \beta_9\text{MUnknown}_{ijkl} + \beta_{10}\text{I1\&2}_{ijkl} + \beta_{11}\text{I3\&4}_{ijkl} + \beta_{12}\text{I7\&8}_{ijkl} + \beta_{13}\text{IUnknown}_{ijkl} + \beta_{16}\text{GDP}_{kl} + \beta_{17}(\text{Wilkinson threshold})_{kl} + \beta_{18}(\text{Wilkinson threshold})_{kl}*(\text{GDP-10})_{kl} + \boldsymbol{\beta_{19}(\text{EHII-40})_{kl} + \beta_{20}(\text{Wilkinson threshold})_{kl}*(\text{EHII-40})_{kl}}$$

$+ (\mu_{01}X_{0ijkl} + \nu_{0kl} X_{0ijkl} + \mu_{141}\text{Low-income}_{ijkl} + \mu_{151}\text{Unknown-income}_{ijkl})$

$E(\log(\pi_{ijkl}^{fair} / \pi_{ijkl}^{good})) = \beta_{0kl}X_{0ijkl} + \beta_1 W2_{ijkl} + \beta_2 W3_{ijkl} + \beta_3 W4_{ijkl} + \beta_4(\text{Age-40})_{ijkl} + \beta_5\text{Female}_{ijkl} + \beta_6(\text{Age-40})*\text{Female}_{ijkl} + \beta_7\text{Msingle}_{ijkl} + \beta_8\text{Mseperated/window/divoced}_{ijkl} + \beta_9\text{MUnknown}_{ijkl} + \beta_{10}\text{I1\&2}_{ijkl} + \beta_{11}\text{I3\&4}_{ijkl} + \beta_{12}\text{I7\&8}_{ijkl} + \beta_{13}\text{IUnknown}_{ijkl} + \beta_{16}\text{GDP}_{kl} + \beta_{17}(\text{Wilkinson threshold})_{kl} + \beta_{18}(\text{Wilkinson threshold})_{kl}*(\text{GDP-10})_{kl} + \boldsymbol{\beta_{19}(\text{EHII-40})_{kl} + \beta_{20}(\text{Wilkinson threshold})_{kl}*(\text{EHII-40})_{kl}} + (\mu_{01}X_{0ijkl} + \nu_{0kl} X_{0ijkl} + \mu_{141}\text{Low-income}_{ijkl} + \mu_{151}\text{Unknown-income}_{ijkl})$

where EHII-40 means the estimated household income inequality centred on the value of 40, The potential range of this variable is from zero for a country with the most equal distribution to one hundred, the most unequal. So here the base category is the group of respondents who are middle income, 40-year-old married men who live in a country with average income of ten thousands US dollars and its estimated household income distribution is 40 in 1981 cohort. The term β_{19} is the degree that increasing one unit of income inequality will affect people's perception of reporting their health condition. The coefficient β_{19}; the interaction between EHII-40 and Wilkinson threshold, estimates the relationship between health and income inequality separately for countries above five thousands US dollars in GDP purchase power parity per capita in 1990. This model allows us to test the relative income hypothesis for all countries as well as countries separately in term of countries above or below Wilkinson's threshold.

Model 5: builds on Model 4 but includes the cross-level interaction of country income inequality and individual income groups. This will give us the not only the relationship between income inequality and self-rated health, but also show the differential impact that income inequality has on self-rated health of individuals who are on different incomes. This is because Wilkinson argues that poor people suffer more in an unequal country than rich people.

The model is specified as follows:

$E(\log(\pi_{ijkl}^{poor} / \pi_{ijkl}^{good}))$ = $\beta_{0kl}X_{0ijkl}$ + $\beta_1 W2_{ijkl}$ + $\beta_2 W3_{ijkl}$ + $\beta_3 W4_{ijkl}$ + $\beta_4(Age\text{-}40)_{ijkl}$ + $\beta_5 Female_{ijkl}$ + $\beta_6(Age\text{-}40)*Female_{ijkl}$ + $\beta_7 Msingle_{ijkl}$ + $\beta_8 Mseperated/window/divoced_{ijkl}$ + $\beta_9 MUnknown_{ijkl}$ + $\beta_{10} I1\&2_{ijkl}$ + $\beta_{11} I3\&4_{ijkl}$ + $\beta_{12} I7\&8_{ijkl}$ + $\beta_{13} IUnknown_{ijkl}$ + $\beta_{16} GDP_{kl}$ + $\beta_{17}(Wilkinson\ threshold)_{kl}$ + $\beta_{18}(Wilkinson\ threshold)_{kl}*(GDP\text{-}10)_{kl}$ + $\beta_{19}(EHII\text{-}40)_{kl}$ + $\beta_{20}(Wilkinson\ threshold)_{kl}*(EHII\text{-}40)_{kl}$ + $\boldsymbol{\beta_{21}(EHII\text{-}40)_{kl}*I1\&2_{ijkl}}$ + $\boldsymbol{\beta_{22}(EHII\text{-}40)_{kl}*I3\&4_{ijkl}}$ + $\boldsymbol{\beta_{23}(EHII\text{-}40)_{kl}*I7\&8_{ijkl}}$ + $\boldsymbol{\beta_{24}(EHII\text{-}40)_{kl}*I9\&10_{ijkl}}$ + $\boldsymbol{\beta_{25}(EHII\text{-}40)_{kl}*IUnknown_{ijkl}}$ + $(\mu_{0l}X_{0ijkl} + \nu_{0kl}X_{0ijkl}$ + $\mu_{14l}Low\text{-}income_{ijkl} + \mu_{15l}Unknown\text{-}income_{ijkl})$

$E(\log(\pi_{ijkl}^{fair} / \pi_{ijkl}^{good}))$ = $\beta_{0kl}X_{0ijkl}$ + $\beta_1 W2_{ijkl}$ + $\beta_2 W3_{ijkl}$ + $\beta_3 W4_{ijkl}$ + $\beta_4(Age\text{-}40)_{ijkl}$ + $\beta_5 Female_{ijkl}$ + $\beta_6(Age\text{-}40)*Female_{ijkl}$ + $\beta_7 Msingle_{ijkl}$ + $\beta_8 Mseperated/window/divoced_{ijkl}$ + $\beta_9 MUnknown_{ijkl}$ + $\beta_{10} I1\&2_{ijkl}$ + $\beta_{11} I3\&4_{ijkl}$ + $\beta_{12} I7\&8_{ijkl}$ + $\beta_{13} IUnknown_{ijkl}$ + $\beta_{16} GDP_{kl}$ + $\beta_{17}(Wilkinson\ threshold)_{kl}$ + $\beta_{18}(Wilkinson\ threshold)_{kl}*(GDP\text{-}10)_{kl}$ + $\beta_{19}(EHII\text{-}40)_{kl}$ + $\beta_{20}(Wilkinson\ threshold)_{kl}*(EHII\text{-}40)_{kl}$ + $\boldsymbol{\beta_{21}(EHII\text{-}40)_{kl}*I1\&2_{ijkl}}$ + $\boldsymbol{\beta_{22}(EHII\text{-}40)_{kl}*I3\&4_{ijkl}}$ + $\boldsymbol{\beta_{23}(EHII\text{-}40)_{kl}*I7\&8_{ijkl}}$ + $\boldsymbol{\beta_{24}(EHII\text{-}40)_{kl}*I9\&10_{ijkl}}$ + $\boldsymbol{\beta_{25}(EHII\text{-}40)_{kl}*IUnknown_{ijkl}}$ + $(\mu_{0l}X_{0ijkl} + \nu_{0kl}X_{0ijkl}$ + $\mu_{14l}Low\text{-}income_{ijkl} + \mu_{15l}Unknown\text{-}income_{ijkl})$

where there are five more cross-level interaction terms between a country's income inequality centred on 40 and each category of individual income except 'I5&6' which has been treated as base. So here the base is the group of respondents of the middle group, 40-year-old married men who live in a country with average income of ten thousands US dollars and its estimated household income distribution is 40 in 1981 cohort. The terms β_{19} to β_{23} show the difference among different level of individual income groups the differences between means over different levels of a country's income inequality. This model was also fitted to assess if the relationships changed above and below a threshold of GDP above five thousand US dollars in1981.

All models are fitted with accompanying sample weights at level 1 to adjust for potential sampling bias in the survey. Weighting is necessary to approximate national population parameters. In each country, the investigators were asked to provide a 4-digit weight variable to correct their sample to reflect national distributions of key variables such as the distribution of age, education, race, and percentage of urban and rural etc, to obtain a nationally representative sample at particular wave in each country. If no weighting was necessary, each case was simply coded as "1.00."

4.4 RESULTS

Table 4.3 gives the results when applying the multinomial four-level model on self-rated health outcome, from the 69 countries across wave 1 (1981) to wave 4 (1999-2000) and the model-fitted diagnostic DIC. Transforming the logit intercept estimates in Models 1a and 1b to an underlying percentage, 19% of people report themselves in poor health and 61% in fair health in the base year 1981. The time variables of different waves are needed to describe the general trend, as all the estimates of the fixed part coefficients are significant at 0.05. In general, the pattern shows a decreasing trend of the logit of the proportion reporting poor and fair health as can be seen in Figure 4.5, people feel healthier than they used to, but we have to be careful not to over-interpret this trend as different countries were sampled at different times, making temporal comparisons very difficult to make.

In the random part there are significant effects (estimates more than two times the standard error) for all variance and covariance terms at country and wave level (Table 4.4, Model 1a and 1b). The variances at level four -- between-country variation -- suggests that there is a considerable difference between countries in the log-odds of both fair versus good and poor versus good. Moreover the positive covariance and correlation (0.68) between both sets of outcomes suggest that differential logits at the country level for fair/good and poor/good are quite closely related. Countries with high proportions of good in comparison to poor, also tend to have high proportions of fair in comparison to poor

Figure 4.5 The trend of reporting proportion of poor/fair health relatively to good health across study period with 95% confidence bands

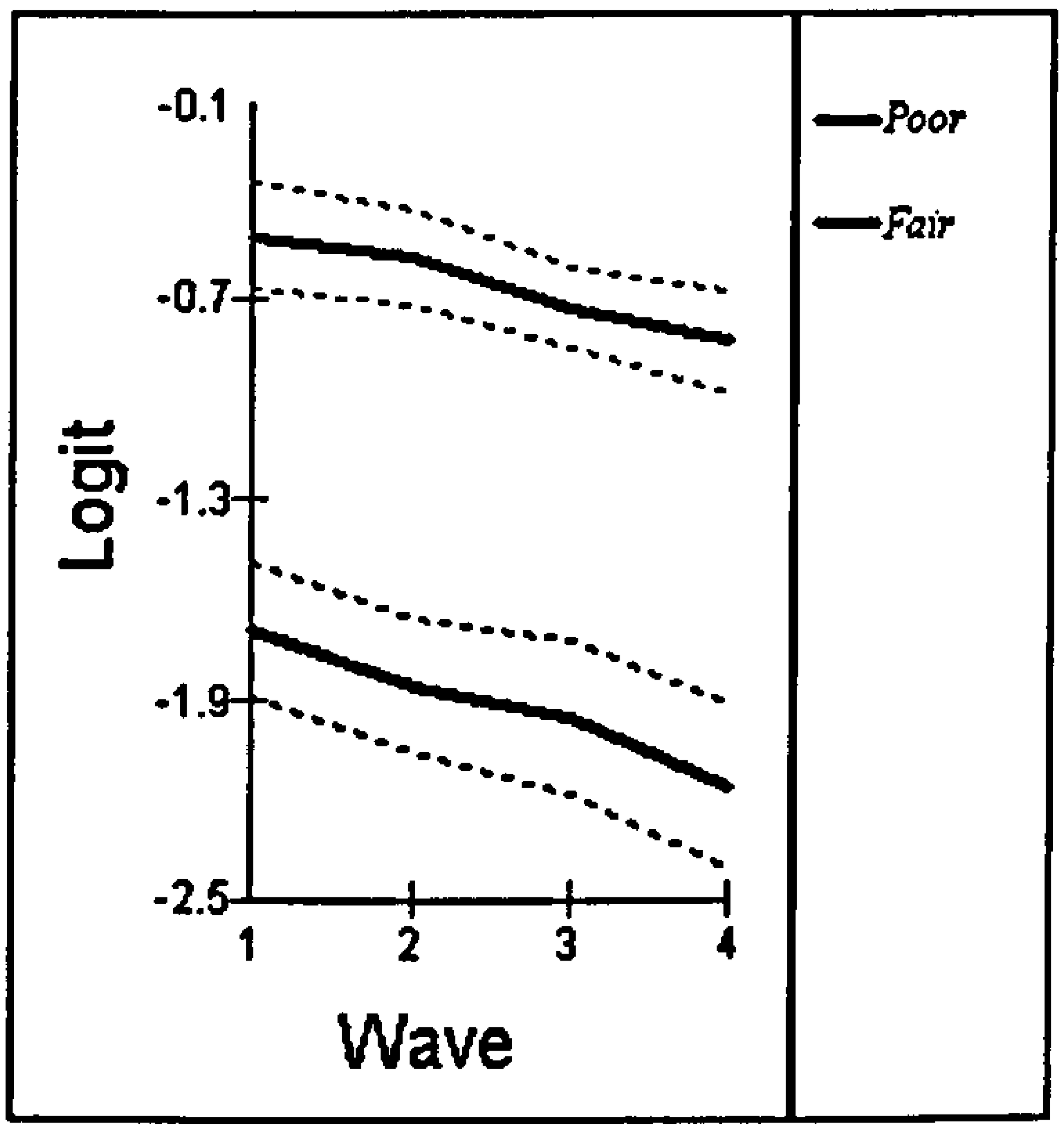


Table 4.3 Results of fixed and random part of the Multinomial analytical models and Deviance information criterion (MCMC); a. refers to Poor/Good and b refers to Fair/Good

Fixed Parameters	Model 1a (S.E.)	Model 1b (S.E.)	Model 2a (S.E.)	Model 2b (S.E.)	Model 3a (S.E.)	Model 3b (S.E.)	Model 4a (S.E.)	Model 4b (S.E.)	Model 5a (S.E.)	Model 5b (S.E.)
Constant	-1.688(0.101)	-0.509(0.084)	-2.217(0.115)	-0.575(0.086)	-2.235(0.245)	-0.447(0.190)	-2.716(0.261)	-0.751(0.184)	-2.726(0.261)	-0.763(0.185)
Wave 1 (base)										
Wave 2	-0.171(0.075)	-0.068(0.039)*	-0.133(0.080)	-0.142(0.034)	0.220(0.128)*	0.075(0.072)*	0.339(0.151)	0.088(0.072)	0.348(0.154)	0.099(0.071)*
Wave 3	-0.263(0.103)	-0.222(0.063)	-0.386(0.096)	-0.353(0.054)	0.012(0.138)*	-0.129(0.089)*	0.401(0.181)	0.055(0.111)*	0.407(0.181)	0.065(0.110)
Wave 4	-0.473(0.132)	-0.325(0.079)	-0.351(0.118)	-0.410(0.074)	-0.020(0.168)*	-0.205(0.113)*	0.259(0.208)*	-0.064(0.139)*	0.275(0.205)*	-0.050(0.138)*
Individual predictors										
Female			0.302(0.037)	0.182(0.017)	0.309(0.038)	0.184(0.018)	0.322(0.043)	0.191(0.020)	0.323(0.043)	0.191(0.020)
Age (centred around 40)			0.043(0.002)	0.023(0.002)	0.042(0.002)	0.023(0.002)	0.045(0.002)	0.024(0.002)	0.045(0.002)	0.024(0.002)
Female*Age			0.000(0.001)*	0.000(0.001)*	0.000(0.001)*	0.001(0.001)*	-0.001(0.002)*	0.000(0.001)*	-0.001(0.002)	0.000(0.001)
Couple (base)										
Widowed/Separate/Divorced			0.248(0.033)	0.029(0.029)*	0.248(0.034)	0.025(0.028)*	0.276(0.037)	0.048(0.027)*	0.275(0.036)	0.049(0.027)*
Single			0.083(0.039)	-0.073(0.037)*	0.076(0.040)*	-0.077(0.037)	0.070(0.047)*	-0.066(0.037)*	0.068(0.047)*	-0.066(0.037)*
Unknown			0.057(0.211)*	0.030(0.154)*	0.068(0.226)*	0.038(0.162)*	0.035(0.239)*	0.024(0.175)*	0.039(0.242)*	0.025(0.175)*
5&6 (base)										
1&2			0.686(0.065)	0.345(0.036)	0.701(0.061)	0.351(0.035)	0.729(0.068)	0.373(0.038)	0.734(0.072)	0.379(0.039)
3&4			0.241(0.058)	0.155(0.026)	0.264(0.056)	0.164(0.025)	0.269(0.065)	0.177(0.028)	0.271(0.064)	0.184(0.029)
7&8			-0.202(0.046)	-0.158(0.023)	-0.196(0.048)	-0.160(0.023)	-0.234(0.053)	-0.174(0.025)	-0.220(0.050)	-0.172(0.026)
9&10			-0.425(0.075)	-0.277(0.036)	-0.406(0.080)	-0.279(0.038)	-0.494(0.087)	-0.291(0.042)	-0.447(0.072)	-0.277(0.041)
Unknown			0.269(0.049)	0.101(0.028)	0.270(0.052)	0.107(0.030)	0.268(0.060)	0.095(0.032)	0.276(0.063)	0.099(0.034)
Country predictors										
GDP (centred around 10K)					0.032(0.026)*	-0.025(0.005)	-0.020(0.029)*	-0.012(0.022)*	-0.022(0.029)*	-0.013(0.022)*
Wilkinson's threshold (base:<5k)					-0.319(0.204)*	-0.417(0.174)	-0.290(0.188)*	-0.327(0.155)	-0.304(0.187)*	-0.333(0.155)
GDP*Wilkinson's threshold					-0.085(0.028)	-0.056(0.021)	-0.046(0.031)*	-0.019(0.023)*	-0.045(0.031)*	-0.019(0.023)*
EHII (centred around 40)							-0.068(0.013)	-0.039(0.011)	-0.071(0.015)	-0.044(0.011)

EHII*Wilkinson's threshold							0.018(0.018)*	0.021(0.013)*	0.020(0.018)*	0.022(0.013)*
Micro-Macro interaction terms										
EHII*1&2									-0.002(0.009)*	0.007(0.005)*
EHII*3&4									-0.006(0.008)*	0.003(0.004)*
EHII*7&8									0.008(0.009)*	0.002(0.003)*
EHII*9&10									0.042(0.009)	0.012(0.007)*
EHII*Unknown									0.007(0.0010)*	0.004(0.006)*
Random parameters										
Level 4	0.445(0.112)	0.268(0.037)	0.599(0.234)	0.369(0.050)	0.501(0.209)	0.296(0.044)	0.619(0.397)*	0.269(0.046)	0.583(0.366)*	0.268(0.047)
Between countries										
Covariance	0.235(0.045)		0.335(0.054)		0.242(0.038)		0.195(0.049)		0.192(0.049)	
High-income			0.599	0.369	0.501	0.296	0.619	0.269	0.583	0.268
Low-income			0.229[3]	0.186	0.181	0.134	0.202	0.122	0.202	0.120
Unknown-income			0.294	0.302	0.222	0.232	0.257	0.219	0.235	0.221
Level 3	0.152(0.037)	0.025(0.006)	0.175(0.046)	0.031(0.008)	0.150(0.040)	0.032(0.007)	0.038(0.010)	0.009(0.002)	0.115(0.030)	0.028(0.006)
Within country between waves										
Covariance	0.006(0.011)*		0.010(0.014)*		0.014(0.013)*		0.002(0.003)*		0.006(0.009)*	
Deviance information criterion (DIC)	269994.20		255946.60		255939.98		255832.45		255810.50	
Change in DIC			14047.60		6.62		107.53		21.95	

All estimates are significant at 0.05 probability level, except those marked by *, which have a probability greater than 0.05.

[3] The variability for a contrasted category depends on the variance of based category, plus variance of contrast category, plus two times covariance of the base and contrast. In this case, the between-country variance of individuals with low income is "$\sigma^2_{\mu 0} + \sigma^2_{\mu 14} + 2 * \sigma_{\mu 014}$" (values from Table 4.4: 0.599+0.286+2*-0.328=0.229), the variance of 'high-income' plus the variance of 'low-income' and plus two times the association between the differential intercept and the differential slope for 'low-income' and vice versa for other groups.

Table 4.4 Random variance matrix at the country level and wave level

At the country level

	Model 1a & 1b	
	μ_{0l}^{poor}	μ_{0l}^{fair}
μ_{0l}^{poor}	0.445 (0.112)	
μ_{0l}^{fair}	0.235 (0.045)	0.268 (0.037)
μ_{14l}^{poor}		
μ_{14l}^{fair}		
μ_{15l}^{poor}		
μ_{15l}^{fair}		

	Model 2a & 2b					
	μ_{0l}^{poor}	μ_{0l}^{fair}	μ_{14l}^{poor}	μ_{14l}^{fair}	μ_{15l}^{poor}	μ_{15l}^{poor}
μ_{0l}^{poor}	0.599 (0.234)					
μ_{0l}^{fair}	0.335 (0.054)	0.369 (0.050)				
μ_{14l}^{poor}	-0.328 (0.167)*	-0.127 (0.043)	0.286 (0.113)			
μ_{14l}^{fair}	-0.214 (0.113)*	-0.049 (0.029)*	0.166 (0.075)	0.123 (0.051)		
μ_{15l}^{poor}	-0.154 (0.026)	-0.132 (0.023)	0.048 (0.022)	0.031 (0.015)	0.080 (0.016)	
μ_{15l}^{fair}	-0.122 (0.037)	-0.055 (0.017)	0.027 (0.026)*	0.046 (0.018)	0.036 (0.010)	0.043 (0.010)

	Model 3a & 3b					
	μ_{0l}^{poor}	μ_{0l}^{fair}	μ_{14l}^{poor}	μ_{14l}^{fair}	μ_{15l}^{poor}	μ_{15l}^{poor}
μ_{0l}^{poor}	0.501 (0.209)					
μ_{0l}^{fair}	0.242 (0.038)	0.296 (0.044)				
μ_{14l}^{poor}	-0.288 (0.147)*	-0.102 (0.034)	0.256 (0.100)			
μ_{14l}^{fair}	-0.197 (0.104)*	-0.045 (0.025)*	0.149 (0.069)	0.115 (0.049)		
μ_{15l}^{poor}	-0.127 (0.021)	-0.119 (0.022)	0.043 (0.019)	0.025 (0.015)*	0.076 (0.016)	
μ_{15l}^{fair}	-0.099 (0.029)	-0.052 (0.017)	0.022 (0.020)*	0.043 (0.016)	0.035 (0.010)	0.041 (0.010)

	Model 4a & 4b					
	μ_{0l}^{poor}	μ_{0l}^{fair}	μ_{14l}^{poor}	μ_{14l}^{fair}	μ_{15l}^{poor}	μ_{15l}^{poor}
μ_{0l}^{poor}	0.620 (0.397)					
μ_{0l}^{fair}	0.195 (0.049)	0.269 (0.046)				
μ_{14l}^{poor}	-0.340 (0.252)*	-0.081 (0.036)	0.263 (0.157)*			
μ_{14l}^{fair}	-0.257 (0.182)*	-0.028 (0.027)*	0.171 (0.113)*	0.151 (0.082)*		
μ_{15l}^{poor}	-0.112 (0.024)	-0.115 (0.028)	0.034 (0.020)*	0.022 (0.016)*	0.083 (0.020)	
μ_{15l}^{fair}	-0.115 (0.043)	-0.048 (0.019)	0.041 (0.028)*	0.060 (0.022)	0.041 (0.012)	0.046 (0.011)

	Model 5a & 5b					
	μ_{0l}^{poor}	μ_{0l}^{fair}	μ_{14l}^{poor}	μ_{14l}^{fair}	μ_{15l}^{poor}	μ_{15l}^{po}
μ_{0l}^{poor}	0.583 (0.366)*					
μ_{0l}^{fair}	0.192 (0.049)	0.268 (0.047)				
μ_{14l}^{poor}	-0.307 (0.224)*	-0.080 (0.035)	0.233 (0.134)*			
μ_{14l}^{fair}	-0.249 (0.173)*	-0.027 (0.027)*	0.159 (0.104)*	0.150 (0.081)*		
μ_{15l}^{poor}	-0.114 (0.025)	-0.115 (0.028)	0.036 (0.020)*	0.023 (0.016)*	0.083 (0.019)	
μ_{15l}^{fair}	-0.116 (0.043)	-0.049 (0.020)	0.040 (0.027)*	0.062 (0.023)	0.042 (0.012)	0.050 (0.012)

14 refers the low-income group, 15 refers the unknown-income group

At the wave level

	Model 1a & 1b		Model 2a & 2b		Model 3a & 3b		Model 4a & 4b		Model 5a & 5b	
	v_{0k}^{poor}	v_{0k}^{fair}	v_{0k}^{poor}	v_{0k}^{fair}	v_{0k}^{poor}	v_{0k}^{fair}	v_{0k}^{poor}	v_{0k}^{fair}	v_{0k}^{poor}	v_{0k}^{fair}
v_{0k}^{poor}	0.152 (0.037)		0.175 (0.046)		0.150 (0.040)		0.038 (0.010)		0.115 (0.030)	
v_{0k}^{fair}	0.006 (0.011)*	0.025 (0.006)	0.010 (0.014)*	0.031 (0.008)	0.014 (0.013)*	0.032 (0.007)	0.002 (0.003)*	0.009 (0.002)	0.006 (0.009)*	0.028 (0.006)

All estimates are significant at 0.05 probability level, except those marked by *, which have a probability greater than 0.05.

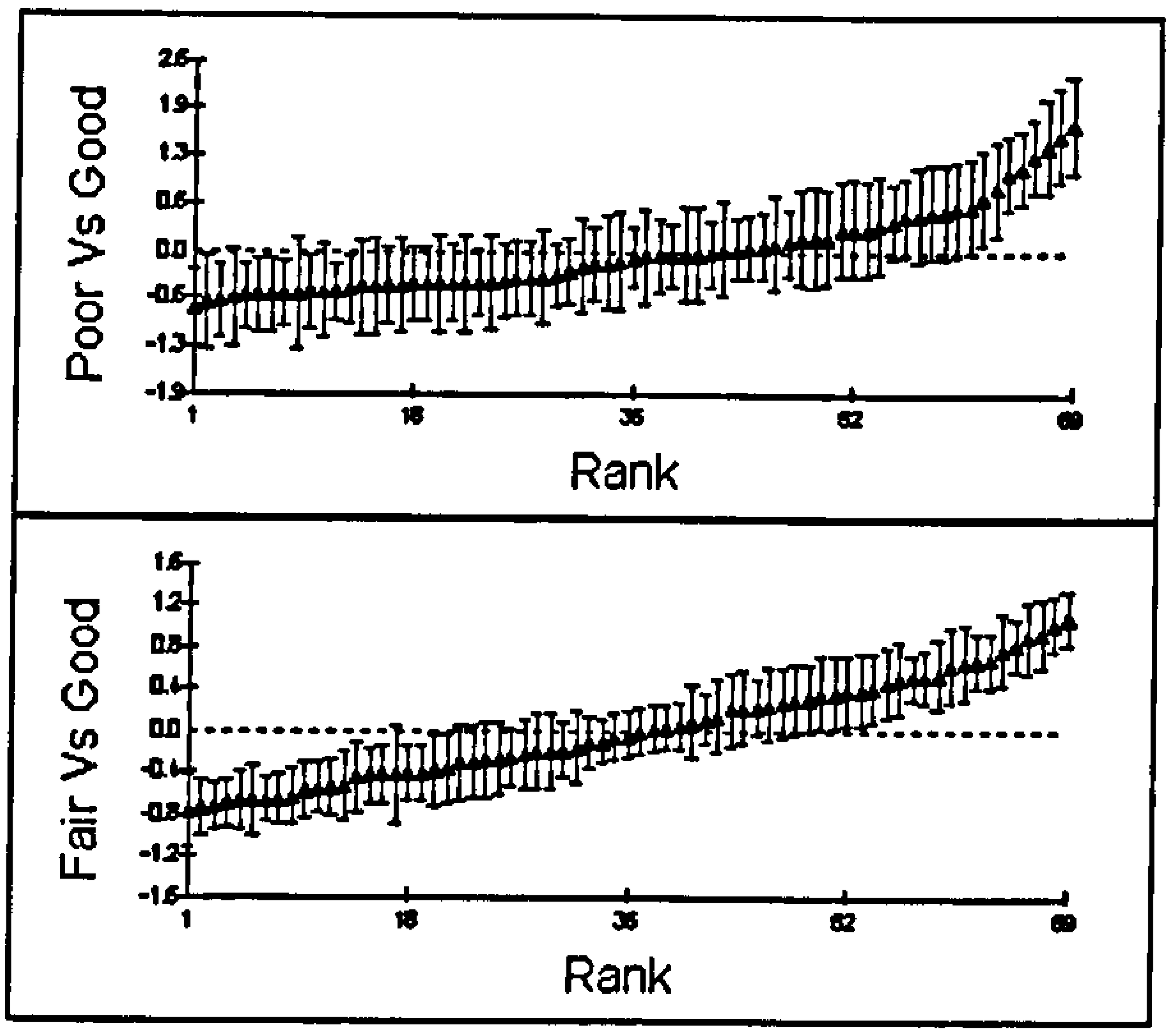
Figure 4.6 shows the between country variation as point estimates on a logit scale, for poor and fair health in comparison to good. These are values derived across all the waves of the survey. In 4.6a), a 'caterpillar plot' shows the logit estimates and their 95% confidence intervals plotted against rank. The horizontal dashed line indicates the average of all countries. Quite a few countries are significantly above and below the average at both ends statistically. Consequently, the average logit proportion reporting poor/fair health in a country is very different from each other.

Table 4.5 shows the fifteen 'best' and fifteen 'worst' estimates of the differentials for countries. Thus the extremes for reporting poor health when transformed to a Standard Morbidity Ratio are Ukraine at 486 and Switzerland at 27, compared to an all-country average of 100 across all waves. There is therefore a 13 fold difference between countries with the most and least poor health. It is noticeable in this listing that nine out of fifteen of worst countries are those for former Communist states; most of the fifteen best countries are dominated by wealthy economies. These are very substantial differences and give our first evidence that countries matter in terms of health perception.

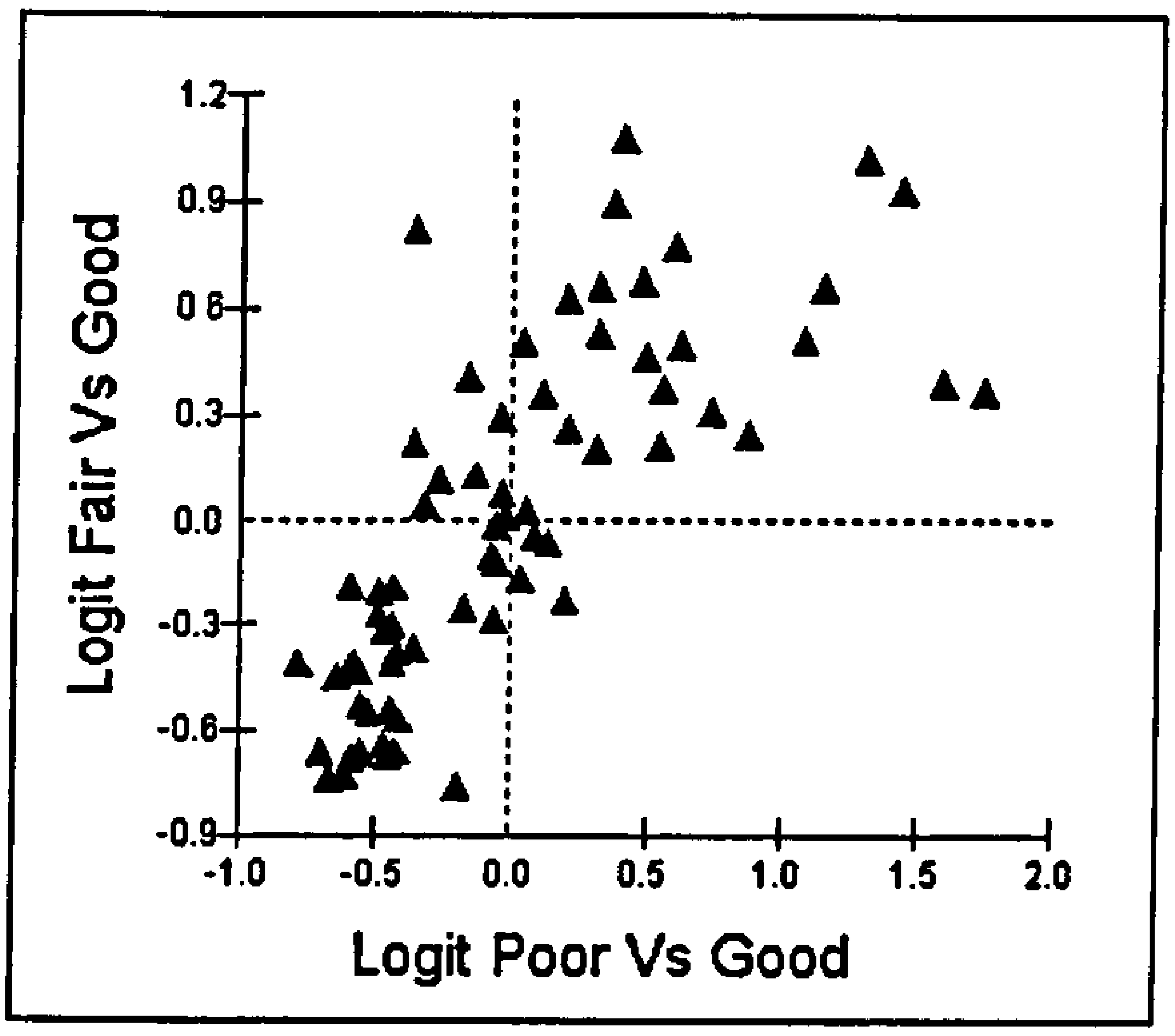
In the scatter plot in 4.5 b) which plots the differential country level logits, there is a tendency for countries with a relatively high amount of poor health to also experience a high degree of fair (as opposed to good) health as would be expected from the positive covariance. Turning now to the level 3 variances and co-variances within countries between waves, it can be seen that these are smaller than between countries. Thus, there was some variation within countries across waves, but the effects are not as substantial as between country differences (see Figure 4.6).

Figure 4.6 Between country differences

a) Logit and 95% confidence intervals against rank of country



b) Scatter plot of the differential country logits for Poor and Fair health

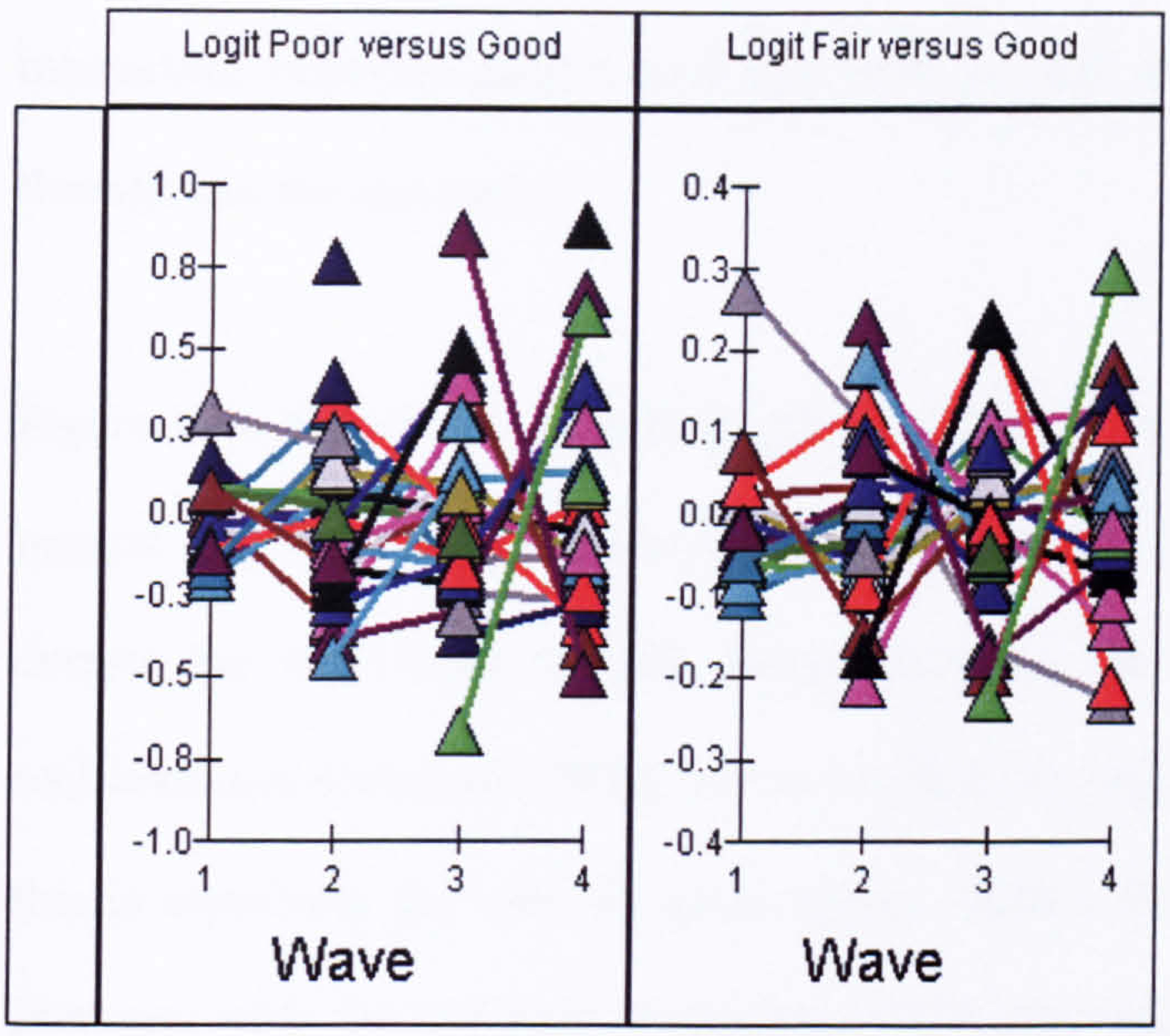


Those countries with poor or fair health tended to remain so across time; although here are some exceptions such as Moldova (coloured purple in the graph) which showed improvement and Bangladesh (light green) which showed an increase in poor health. Overall, there are one or two bigger changes around generally small changes. Notice that it is not the same vertical scale for both plots, so that Fair versus good is much less volatile than Poor versus good.

Table 4.5 Country differentials in self-rated health; fifteen worst and fifteen best countries in terms of difference in 1981 of reporting poor/good and fair/good health of in logit and odds

(Poor/Good) Countries	Differential from global average in 1981 in Logit	Differential from global average in 1981 in odds	Rank	(Fair/Good) Countries	Differential from global average in 1981 in Logit	Differential from global average in 1981 in odds
Ukraine	1.581	486	1	Russia	1.422	414
Russia	1.556	474	2	Ukraine	1.377	396
Hungary	1.371	394	3	Latvia	1.164	320
Jordan	1.239	345	4	Moldova	0.920	251
Moldova	1.137	312	5	Peru	0.901	246
Bangladesh	1.059	288	6	Estonia	0.884	242
Poland	0.814	226	7	Algeria	0.879	241
Estonia	0.782	218	8	Hungary	0.845	233
Slovenia	0.778	218	9	Bangladesh	0.831	230
Latvia	0.765	215	10	Slovakia	0.713	204
Belarus	0.688	199	11	Jordan	0.674	196
Lithuania	0.687	199	12	Dominic Rep.	0.630	188
Portugal	0.608	184	13	Lithuania	0.628	187
Georgia	0.566	176	14	Poland	0.626	187
Algeria	0.547	173	15	Taiwan	0.578	178
……	……		……	……	……	
Australia	-0.843	43	55	Netherlands	-0.582	56
Brazil	-0.853	43	56	Uruguay	-0.638	53
Denmark	-0.894	41	57	New Zealand	-0.725	48
Ghana	-0.904	41	58	S Africa	-0.725	48
Sweden	-0.934	39	59	Sweden	-0.730	48
New Zealand	-0.941	39	60	Britain	-0.746	47
Belgium	-0.965	38	61	Australia	-0.856	42
USA	-1.009	36	62	Norway	-0.878	42
Netherlands	-1.062	35	63	USA	-0.898	41
Canada	-1.159	31	64	Denmark	-0.921	40
Uruguay	-1.176	31	65	Switzerland	-0.963	38
Nigeria	-1.218	30	66	Nigeria	-0.964	38
Ireland	-1.247	29	67	Canada	-1.015	36
S Korea	-1.253	29	68	Ireland	-1.038	35
Switzerland	-1.300	27	69	Indonesia	-1.129	32

Figure 4.7 The residual of different countries across waves



* Note these are differential over time, not the country difference per se (because this is within country between wave differences; that is they are level 3 residuals

Model 2 additionally includes a range of individual variables for respondents, representing social and demographic characteristics. Many of the effects are significant and substantial with a 14047.6 decrease in DIC from 269994.20 in Model 1 to 255946.60 in Model 2 (see Table 4.3). A series of graphs showing the effects of individual characteristics are given in Figure 4.8. These are given in terms of odds (which have been transformed from the logits) in comparison to the base category of middle-income 40-year-old married men in 1981.

Figure 4.8a shows the individual effect of age and sex. As expected, the elderly tend to report greater poor and fair health in comparison to good than the young people, and women are more likely than men to report themselves in poor and fair health. In addition to the reproductive processes, there is a wide range of genetic, hormonal, and metabolic influences playing a part in shaping distinctive male and female patterns of morbidity and mortality (Doyal, 2001). It is also the case, reflected in this study that women tend to report more morbidity than men. However, there is no significant interaction between gender and age here so that the excess for women is found through out the age range.

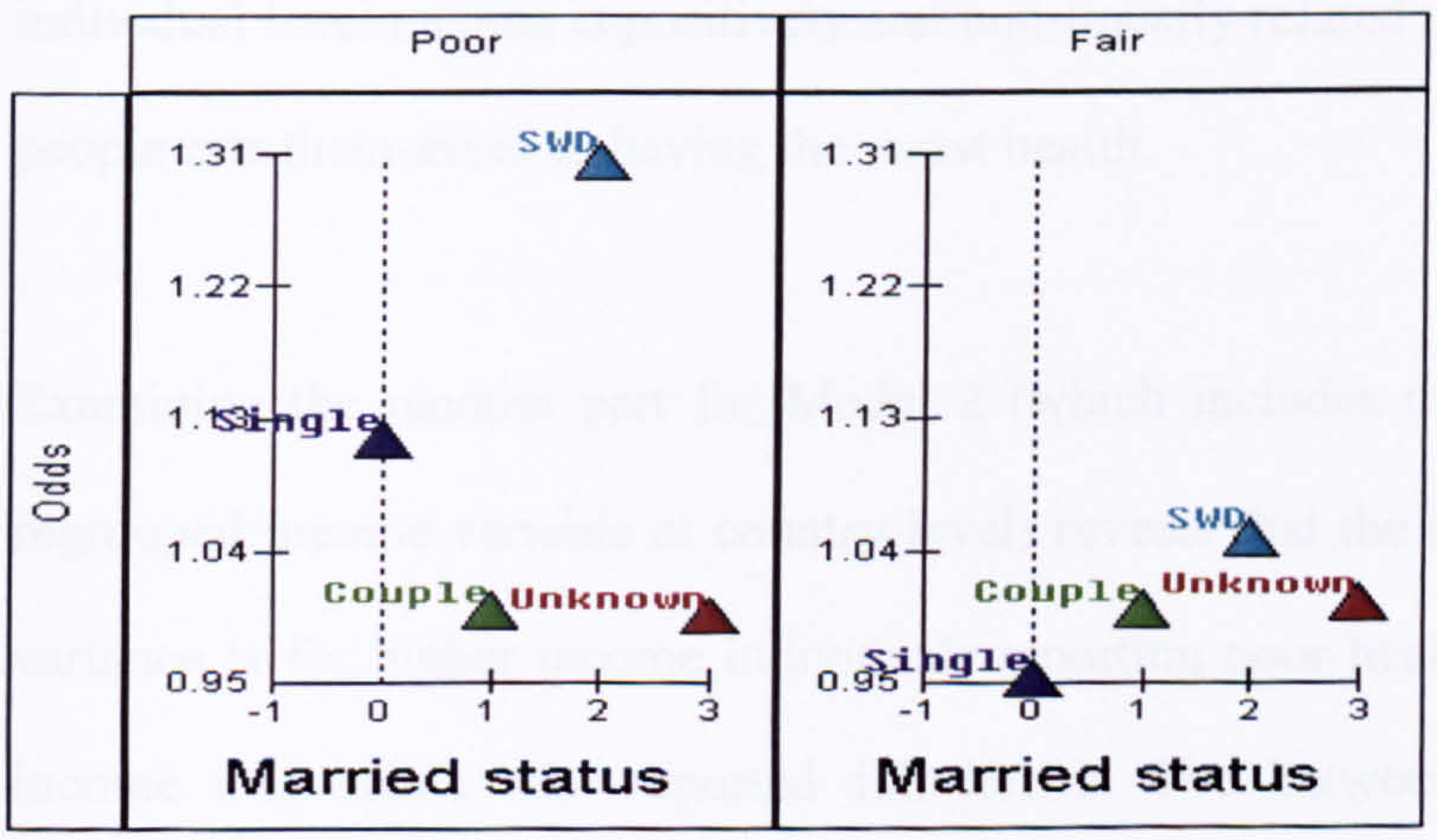
Figure 4.8b shows the individual effect of marital status. None of the categories of marital status is significantly different from each other for reporting fair health, despite the very large sample. Respondents in the marital situation of separated, widowed and divorced (SWD), are more likely to report poor/fair health than couples; this is especially the case for poor health. Such a finding is supported in the wider literature with, for example, Gallagher (1999) showing that all parties to a divorce are likely to suffer increased morbidity and mortality. Those in the Single category experience lower odds than Couples for reporting fair health (but not significant at $p>0.05$), but higher odds in saying that they are in poor health (significant at $p<0.05$); Where marital status is Unknown, there are the essentially same odds of reporting poor and fair health as Couples.

The terms that we have so far discussed, while interesting in their own right, have mainly been included as 'controls'. However Figure 4.8c shows the relation with income group which is a key part of our assessment of the Wilkinson hypothesis. It shows a very consistent 'dose-response' relationship such that the more income a person has the healthier they are. Apart from the Unknown group, increasing odds of poor and fair health are reported as income goes from the wealthiest quintile (9&10) to the lowest quintile (1&2). Every single differential income group is significant from each other at 95% confidence interval for both reporting poor and fair health. There is almost a four-fold difference for reporting poor health, and a two-fold difference for fair health between the top and bottom income groups.

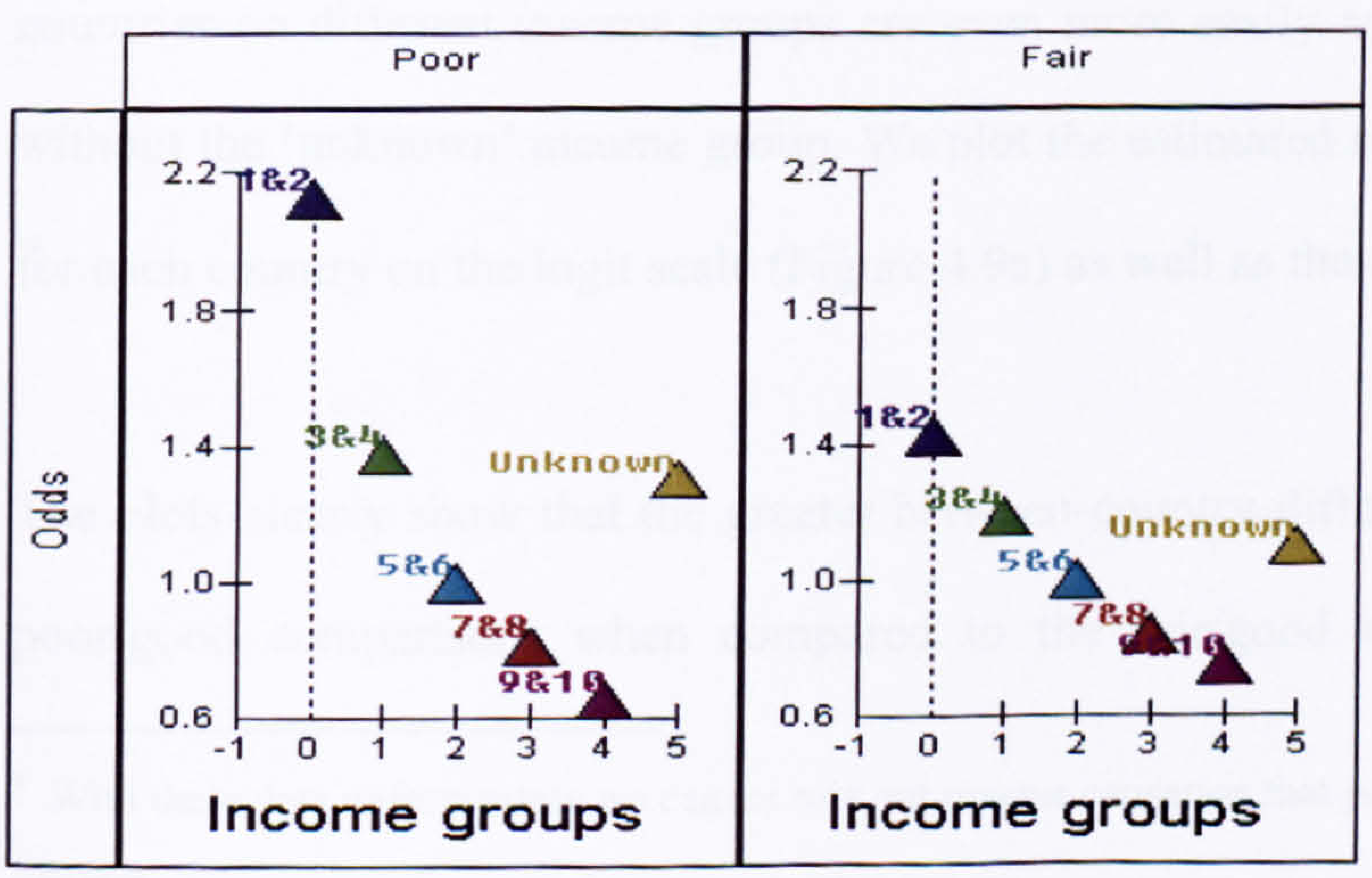
Figure 4.8 individual effects

a) Age and Sex

b) Marital status



c) Income groups



These are the first results to be produced on a global scale, but they accord with a growing literature from national studies that individual earnings has a large effect on all of the health pathways such as morbidity, depressive symptoms and follow-up self-rated health. For example, Fiscella and Franks (2000) concluded that individual income has a large effect on self-rated health and mortality in the USA[4]. Crucially, the relationship between self-reported health and individual income is markedly non-linear in the odds, with the difference between the income groups increasing at higher incomes. Moreover, this is not just an artefact of using odds because this non-linear pattern is to be found in the logits for both reporting poor and fair health. This means that the first part of Gravelle's (1998) supposition is supported: at the individual level income is positively and non-linearly related to self-rated health; poor people rate themselves as having the worst health.

Examining the random part for Model 2 (which includes the random slope of the regrouped income variable at country-level) reveals that the largest between-country variance is for higher income individuals reporting poor health but smallest for low income individuals who reported fair health. The between-country variances of different income categories are shown in Table 4.3. The degree of variation between countries on different income groups are even more easily appreciated in Figure 4.9 without the 'unknown' income group. We plot the estimated relationship with income for each country on the logit scale (Figure 4.9a) as well as the odds (Figure 4.9b).

The plots clearly show that the greater between-country difference are found for the poor/good comparisons when compared to the fair/good relation and that High

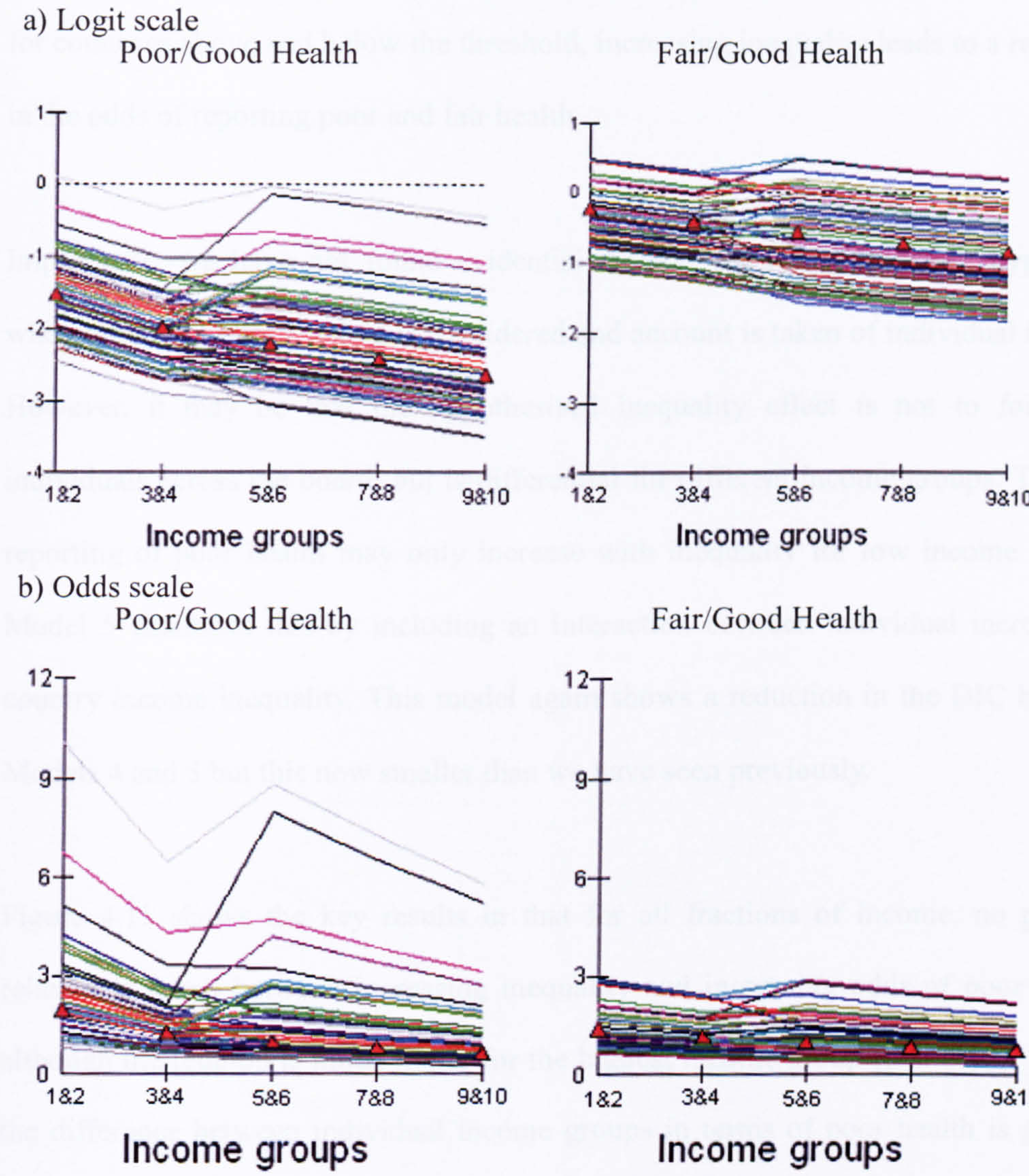
[4] With these data unfortunately we cannot rule out reverse causation that poor health produces lower income.

income groups report less bad health. However, there are some exceptions, for example, Belgium (the grey line) and Jordan (black line) where the high-income people from these two countries have a higher rate of reporting poor health versus good health. Spain has the shallowest relationship within individual income. Importantly, there is still significant and substantial between-country variations even after gender, age and individual income and marital status, have been taken into account. Thus, all the remaining variation is not only at the individual level, but also the country-context has an effect.

The next model, Model 3, includes the country by wave variable GDP, Wilkinson's threshold which is a categorical variable classifying countries above or below five thousand dollars in their GDP in 1990, and their interaction to investigate not only the effect of GDP on self-rated health for all countries but also the differences between countries of above or below Wilkinson's threshold. This model is an improvement from Model 2 with a reduction in the DIC of 6.62.

In Model 4, wave-specific country income inequality is additionally included in the model to look at its effect on health after taking account income effect of individuals as well as countries. The effect of inequality is also allowed to be different in the advanced economies above the Wilkinson's threshold. However, there are 8 countries (Estonia, Georgia, Dominic Republic, Belarus, Slovakia, VietNam, Uganda, United Republic of Tanzania) which do not have the information for the inequality index and these will be excluded in Models 4 and 5. They are countries with a relatively smaller sample size and the results of only 61 countries are similar in terms of Model 3 to those with all 69 countries included in the model. Again this model is an improvement on the simpler Model 3 with a reduction in the DIC of over 100 for comparable models.

Figure 4.9 Residuals of different income groups in the country level after taking account of individual social demographic variables



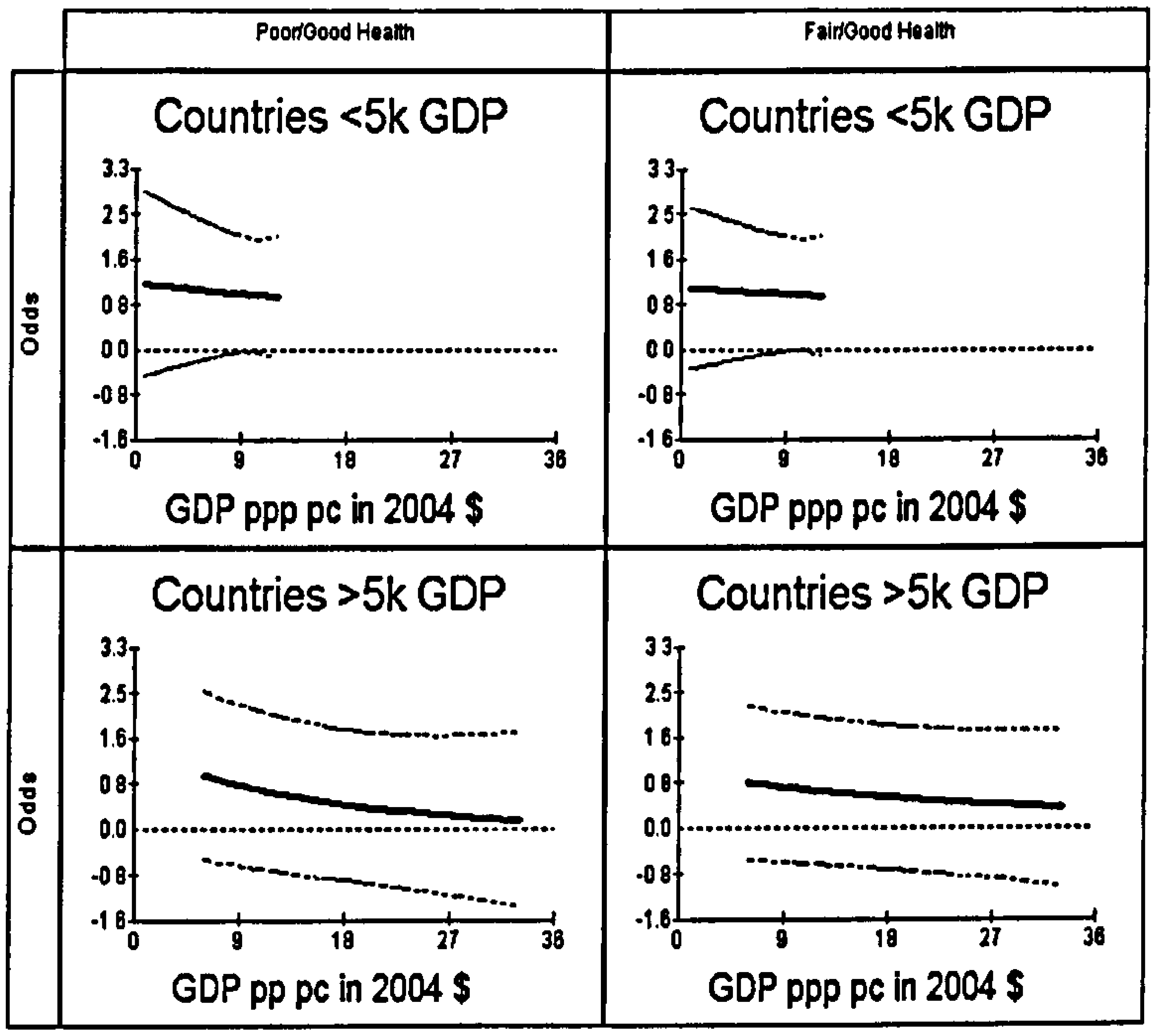
The key results from Model 4 for the Wilkinson hypothesis are plotted in Figure 4.10. The plots show the effect of GDP and Inequality on the odds of reporting poor and fair health for countries above and below the threshold as well as 95% empirical confidence bands. The relations are after taking account of individual variables. The vertical axes on all the graphs are chosen to have the same scale, so as to make them directly comparable. The effect of GDP on health is found to be negative in both 'rich' and 'poor' countries but the estimated terms are not 'significant'. Moreover for inequality, the results are the exact opposite of that anticipated by Wilkinson in that for countries above and below the threshold, increasing inequality leads to a *reduction* in the odds of reporting poor and fair health.

Importantly, we have not found evidential support for the Wilkinson's hypothesis when individuals are a whole is considered and account is taken of individual income. However, it may be that the hypothesised inequality effect is not to found for individuals across the board, but is differential for different income groups. Thus the reporting of poor health may only increase with inequality for low income groups. Model 5 examines this by including an interaction between individual income and country income inequality. This model again shows a reduction in the DIC between Models 4 and 5 but this now smaller than we have seen previously.

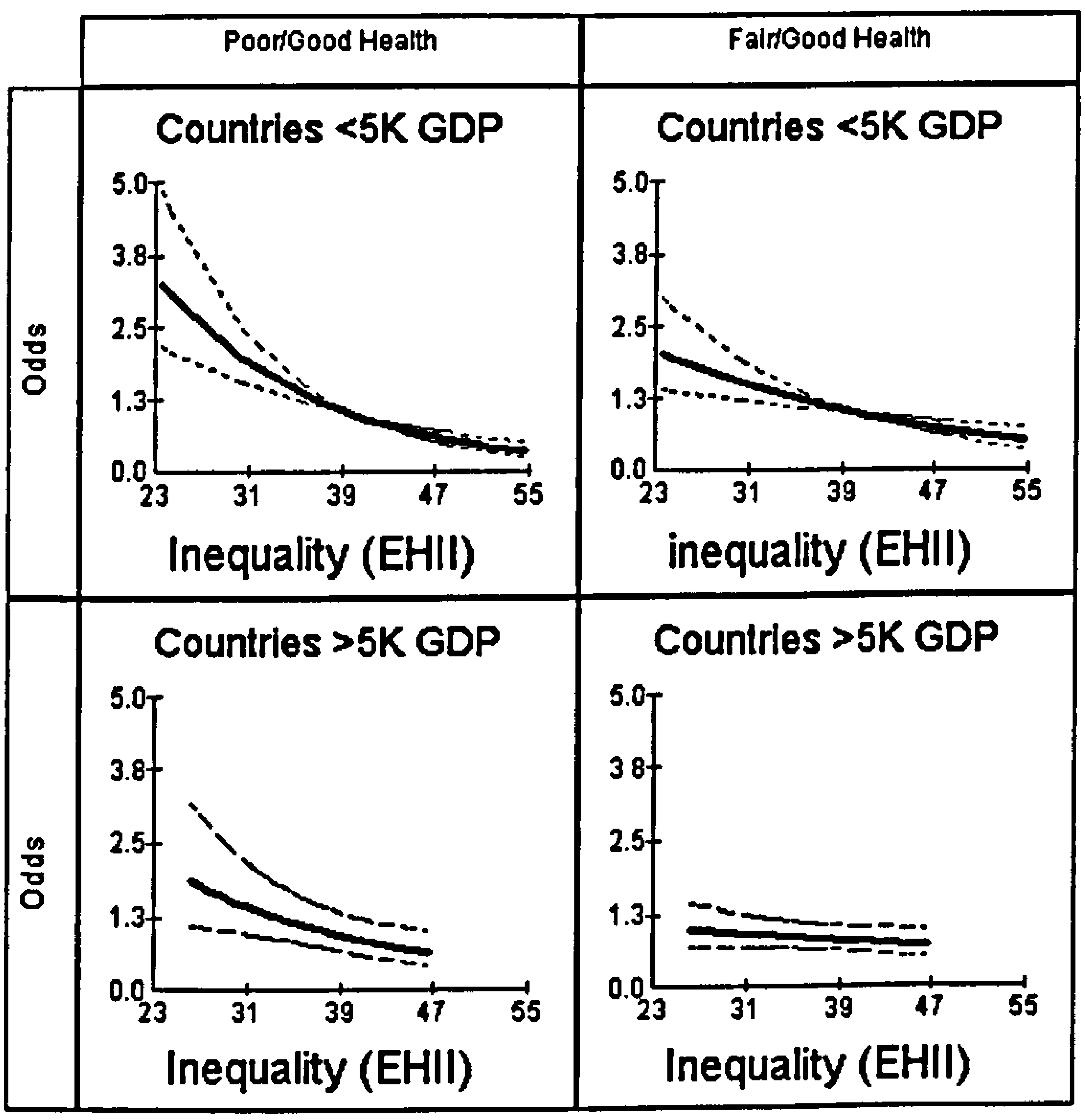
Figure 4.11 shows the key results in that for all fractions of income, no positive relation is found between increasing inequality and increasing odds of poor health, although the relation is much flatter for the highest income group (9 and 10). Indeed, the difference between individual income groups in terms of poor health is greatest when inequality is low. This does not correspond to Wilkinson's statement that in the most unequal countries the poor suffer more than the rich. In summary, there is no support for the relative income hypothesis in terms of self-rated health and this remains the case when the inequality by individual income interaction is allowed to be differentiate above and below the threshold.

Figure 4.10 GDP and Inequalities (EHII) for countries above and below 5k GDP PPP pc in 1990 separately

a) GDP effect

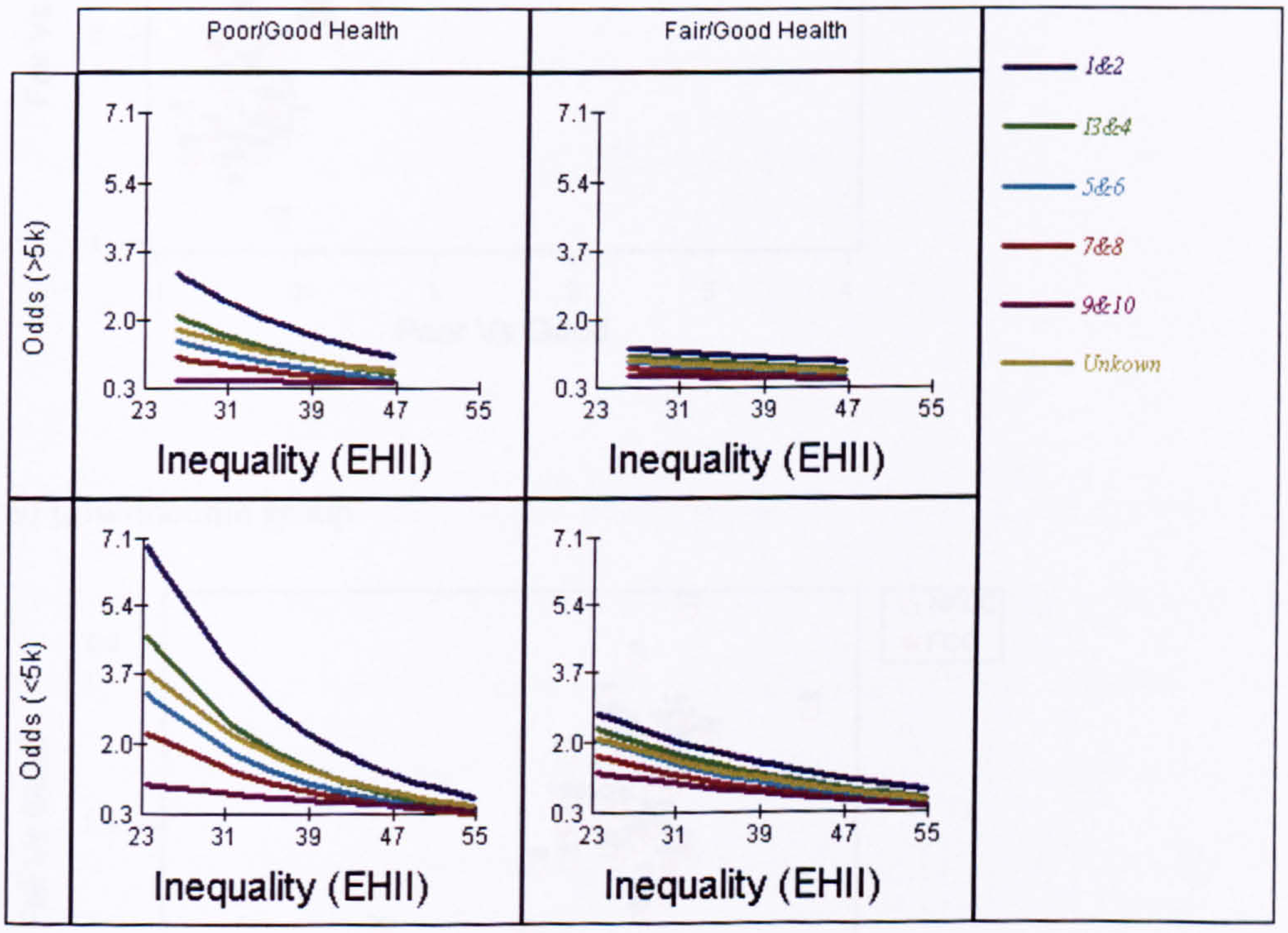


b) EHII effect



Turning now to the between-country random part of the models (Tables 4.3 and 4.4). The striking things is that even after taking account of country GDP and inequality and individual income, as we move from Models 1 to 5, there remains substantial and significant between-country variation that is not explained by the individual and country variables that have been included in the model. Indeed, after taking account of the compositional and contextual variables, the between-country variance is estimated to be somewhat larger. Countries are significantly different in their perceived health status and this particularly true for poor health and for the higher income groups.

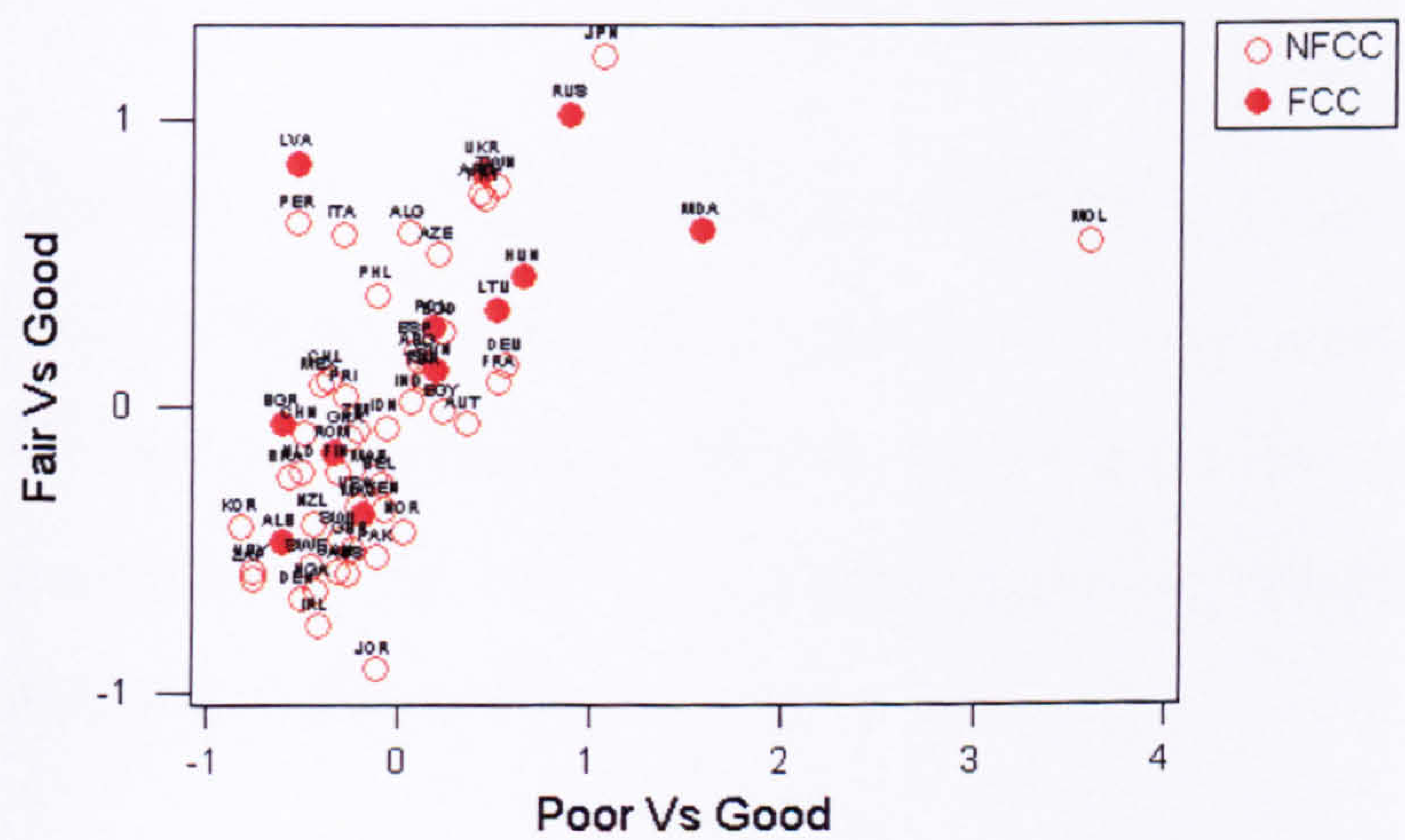
Figure 4.11 The differential patterns of reporting not good health for different individuals' income in different income inequality countries



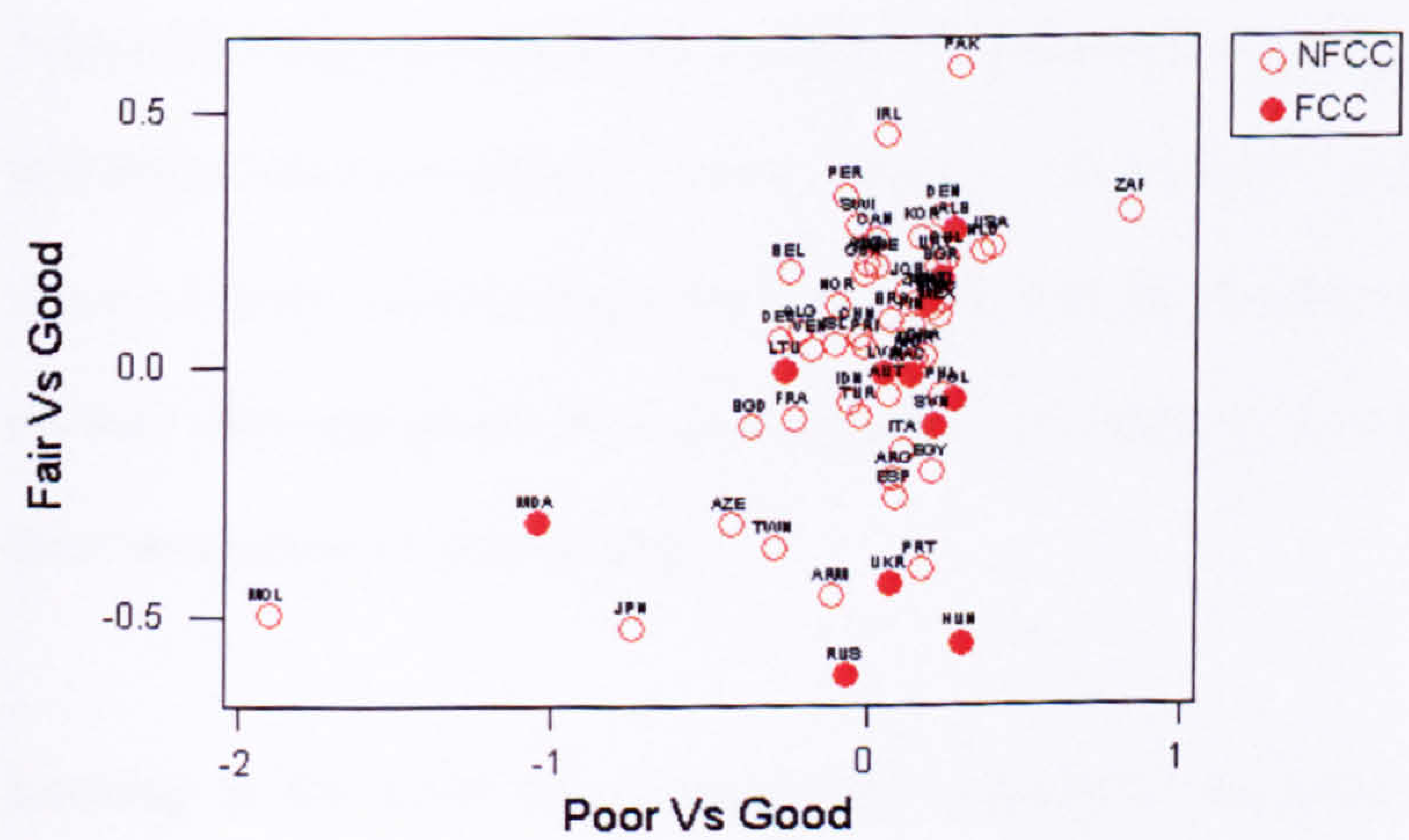
The between-country variations are most clearly seen in a set of residuals plots (Figure 4.12) which graph the logit estimates for poor and fair health for different individual income groups (high-income, low-income, unknown-income) after taking into account both individual and country variables. There is, however, a suggestive pattern to the results as shown by Figure 4.12a). Many countries belonging to the Former Soviet Bloc have particularly high levels of self-rated poor health; five out of twelve of the poorest health countries are from former communist countries (FCC); they are Hungary, Russia, Ukraine, Poland, and Slovenia.

Figure 4.12 The remaining between-country variation (base on Model 5)

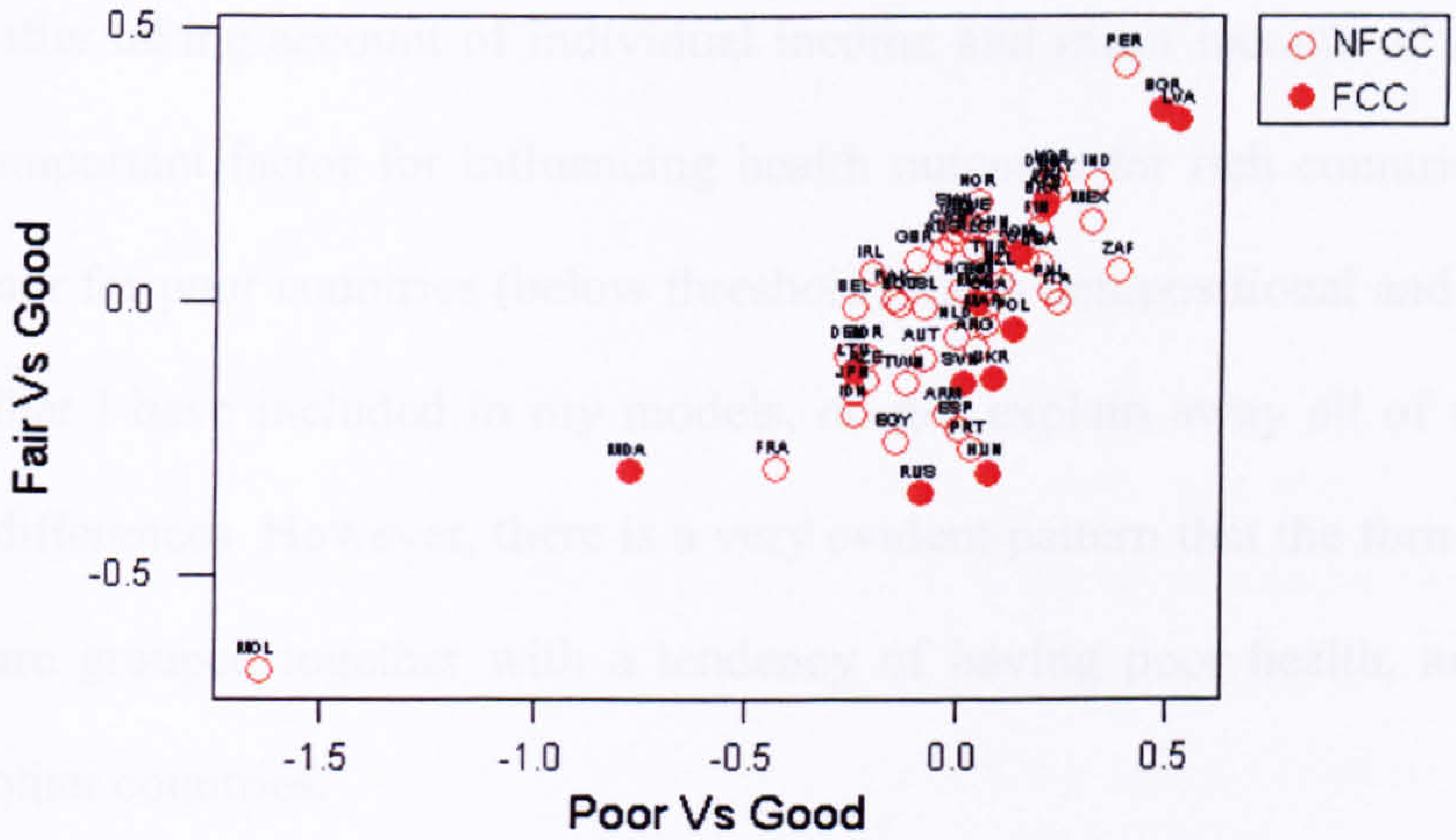
a) High-income group



b) Low-income group



c) Unknown-income group



Moreover, none of the ten best countries included any former communist countries. When comparing between different individuals' income groups, the distinctive pattern between former communist countries and other countries is more obvious in the low-income (Figure 4.12 b) and unknown-income (Figure 4.12 c) groups than high-income (Figure 4.12 a).

4.5 CONCLUSIONS

In this global study of individual's health using the WVS data, I found that there are substantial differences between countries. In terms of the absolute income hypothesis, individual income has been found to have a 'non-linear' 'dose-response' effect with those on lower income reporting more poor and fair health, after taking into account of their demographical characteristics. So the higher a person's wage, the better is their perception of their health.

Looking at the GDP of a country, the relation is the same above and below the threshold of development; GDP has a negative effect but this is only significant for reporting fair health for countries above threshold. The income inequality of a country after taking account of individual income and mean income of the country is not an important factor for influencing health outcome for rich countries (above threshold) nor for poor countries (below threshold). Both compositional and contextual variables that I have included in my models, do not explain away all of the between country differences. However, there is a very evident pattern that the former communist states are grouped together with a tendency of having poor health, and are distinct from other countries.

In my analysis, using a multilevel modelling approach provides the capability to separate out the so called artefact effect, which is introduced by Gravelle, from the true hazarded effect of income inequality. Once account is taken of individual income and its differential impact on self-rated health, there is a lack of support for the Wilkinson hypothesis. GDP has little effect in both rich and poor countries and income inequality doe not have the postulated effect on morbidity. Even when interactions are allowed between individual income and inequality, poor people in the most unequal counties do not appear to experience the poor health predicted by Wilkinson.

Chapter 5

Self-reported health, happiness, and life satisfaction: a global analysis

5.1 INTRODUCTION

The previous chapter was solely concerned with one measure of well-being, that is, self-rated health. While this is a valid instrument for assessing mortality risk, independent of other medical, behavioural and psychosocial risk factors, it does not necessarily capture the multifaceted nature of well-being. According to the definition from the World Health Organisation (2001), health is multidimensional, reflecting dimensions of physical, mental, and subjective well-being (SWB). The nature of the relationship between physical health and subjective well-being is that when our physical health is acutely or chronically compromised, the subjective well-being is reduced (The Royal Society, 2004). Happiness, life satisfaction as well as self-rated health are among vital components of subjective well-being. In this context, they have been studied separately and extensively (such as Kennedy, et al., 1998; Schyns, 2002), across various social and psychological sciences.

As discussed in Chapter 2, Wilkinson considers these three variables as consequences of income inequality. He contends that the key mechanisms between income distribution and health operate through psycho-social processes. Countries where there is greater inequality should experience greater unhappiness, poorer life satisfaction and consequent poorer health. He argues that in developed economies, income inequality operates on health in only a minor way through poor material circumstances (such as inadequate diet, damp housing). For him the key link is that unequal societies are also socially divided societies: this places individuals within such societies in a stressful situation, which then manifests as a range of 'social' problems, including violence, health inequalities and lower life expectancy.

The role of these three variables, happiness, life satisfaction and self-reported health, has been inconsistently investigated. For instant, life satisfaction has been used as an outcome variable by Brief and his colleagues (1993) to be predicted by self-rated health, but in other studies, self-rated health has been specified as the outcome variable while happiness has been consider as a predictor variable (Ostir, *et al.*, 2001; Kubzansky, *et al.*, 2001). However, in this chapter, rather than studying them separately or specifying one component as an outcome and the other as an exposure or a mediator, they have been treated as multiple outcomes for the same individual. Consequently the major aim of this part of the study is to examine the relationships between income and these subjective measures of well-being, at both the individual and country level, and to do so simultaneously. More specifically, I aim to analyse the World Values Survey to show the degree of correlation among self-rated health (on a five-point scale), happiness (on a four-point scale) and life satisfaction (on a ten-point scale), both at an individual level as well as a country level, taking account of wave, age and sex. I intend to develop these models by including individual and country income variables, to further examine the Wilkinson hypothesis that income inequality results in poor life indicators on a range of variables.

In particular, the following research questions will be examined so as to develop the evaluation of the relative income hypothesis:

- Are there differences between countries as well as between people? Do countries that have high self-rated health also have high life satisfaction and happiness?

- How has each of the response variables changed over time?

- What are the relationships between the three dependent variables and individual demographic predictors (age and sex) and income (quintiles)?

- How do the country differences relate to the income variable at the country level?

- Does income inequality (EHII) have an effect over and above individual income and country GDP?

- Does income and income inequality (EHII) have a differential effect above and below the threshold of $5k US dollars in 1990? Is there any difference between countries above and below this threshold?

- Does income inequality (EHII) have an effect over and above individual income and country GDP that is different for different income groups?

The rest of the chapter is organized into three sections: data, methodology and results along with conclusions. The data section starts with description of the chosen dependent variables and their possible determinants in the light of the requirements of testing the relative income hypothesis. Secondly, a *multivariate* multilevel modelling technique is discussed as this is needed to analyse the multiple outcomes for respondents nested in waves and countries. Finally, the results will discuss the correlations between the dependent variables, and the extent to which the correlations are different at the individual, wave and country level, and the effect of income and income inequality on all three dimensions of well-being.

5.2 DATA

5.2.1 The World Value Survey data

As discussed previously, we need individual measurements of the health outcome variables as well as individual demographical and social economic characteristics to evaluate Wilkinson's hypothesis. Continuing to work on the WVS data set, combined with income and income inequality data, we have a number of variables that will allow a direct test of the relative income hypothesis. Three dependent variables from the WVS will be used:

- Self-reported health (variable 11 in the codebook; Inglehart *et al.*, 2000)
- Happiness (variable 10)
- Life satisfaction (variable 65)

Self-rated health is the same variable we used previously and requires no further discussion, Happiness has been assessed on a four point scale by asking the question "Taking all things together, would you say you are: very happy, quite happy, not very happy or not at all happy?". Happiness can be seen as an indication that a person is biologically fit (near to the optimal state) and cognitively in control (capable of counteracting eventual deviations from that optimal state), in other words that he or she can satisfy all basic needs, in spite of possible perturbations from the environment (Heylighen, 1992). According to Veenhoven (1991), life satisfaction is conceived of as "the degree to which an individual judges the overall quality of his life-as-a-whole favourably". In this WVS survey, it is measured by simply asking people directly "All things considered, how satisfied are you with your life as a whole these days?" The responses are given on a ten-point scales with those that the most dissatisfied giving a score of 1, wile the most satisfied giving a score of ten. Appendix 1 gives the exact questions in their English form used in the survey that have been selected in the present study.

I will also use the same individual predictors, (age, sex, marital status, and individual income employed in Chapter Four.

5.2.2 Income and income inequality of countries

This part of the study again uses two global data sets. Country income is measured as GDP per capita (in purchase power parity for 2004 US dollars) on an annual basis from 1970 to 2000 for some 232 countries. Income inequality data comes from the UTIP-UNIDO project at the University of Texas. This has some 3,200 observations over 36 years (1963-1999). Galbraith and Kum (2004) have undertaken a number of qualitative comparisons to confirm that their predictions are an improvement on the lower-quality D&S data (see detail in Chapter Three). Detailed information of both data sets has been provided in the data section of Chapter Three.

The structure is now the multiple responses at level one nested within individuals at level two (the real level one) nested within waves at level three and nested within countries at level four. There are some 69 countries include some 171,214 individuals and slightly over half a million responses (503,699). The key feature of this data set is that the data are available for individuals and have not been aggregated. Importantly, a single survey instrument has been used to collect the data on a large number of countries according to scientific sampling procedures. Consequently, these data provide an ideal empirical test of the relative-income hypothesis as it applies to a range of measures of subjective well-being. Table 5.1 lists the countries and indicates for each wave, whether a country has been surveyed in that specific wave. These data have a complex structure (multivariate repeated cross-sectional) with imbalance where some countries at particular wave did not participate in the survey, and within each wave there are individuals who did not answer all three question (see Figure 5.1).

The response variables exhibit a differing number of missing values, raging from Life satisfaction (4,019) to Happiness (2,896). Missingness (missing data) is ubiquitous in social science research. However, it is very important to discern whether the analyses are valid and under what conditions with missing data. The key is to examine the process by which data have become missing. Rubin (1976) distinguishes data as 'missing at random' (MAR), from data 'missing completely at random' (MCAR), or neither. For data to be missing *completely* at random, the probability that a particular value of a variable is missing is unrelated to the true value of that observation. The data can be considered as missing at random if the data meet the requirement that missingness does not depend on the value of the unobserved value *after controlling for variables*. For example, people who are depressed might be less inclined to report their income. In a model without income as a predictor variable, the missingness would be problematic, while conditioning on income, the missingness is MAR and much less problematic. If the data is MAR then random effects methods such as used in multilevel models, can proceed without explicitly modelling the dropout mechanism (Goldstein and Woodhouse, 1996). In the analysis that follows I have assumed that the data on the responses are missing at random, conditional on demographic and socio-economic characteristics. If this is the case, the random effects multilevel model will not be unduly troubled, (it is just another form of imbalance) and will provide valid estimates of the underlying relationships between the variables.

5.2.3 Preparing to model

The discrete dependent variables: 'self-rated health' has been measured with a five-point scale (poor, fair, good, very good and excellent), 'happiness' with a four-point scale (not at all happy, not very happy, quite happy and very happy) and 'life satisfaction' with a 10-point scale (most unsatisfied to most satisfied). To facilitate analysis these have been replaced with an equivalently-ranked standard normal deviates, so that each has a mean of zero and a standard deviation of 1 (see Table 5.2). The minimum (around -4) and maximum (around 4) show a very symmetric distribution for each variable. This transformation is beneficial for two reasons. First, data have been transformed into a comparable scale among these outcome variables so that we can compare the effects of income on a comparable scale. Second, a multivariate multilevel modelling is based on a multivariate normal distribution assumption (Rasbash *et al.*, 2004) and this assumption will be fulfilled perfectly once the data have been transformed in this way.

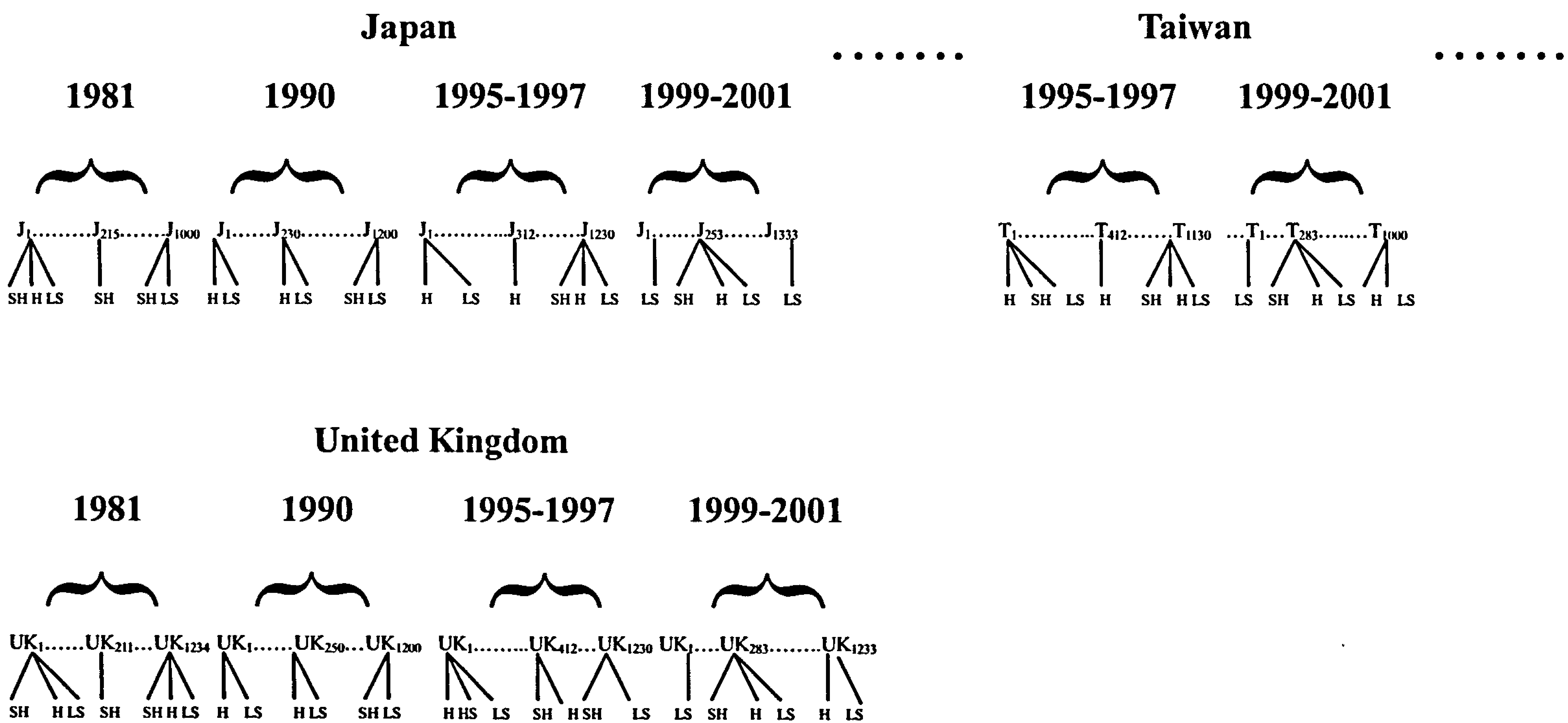
Table 5.1 The structure of the World values Surveys

Code	Country	W1	W2	W3	W4		Code	Country	W1	W2	W3	W4
930	Albania				*		14	Mexico		*	*	*
920	Algeria				*		504	Morocco				*
22	Argentina		*	*	*		5	Netherlands	*	*		*
63	Armenia			*			554	New Zealand			*	
17	Australia	*		*			29	Nigeria		*	*	*
42	Austria		*		*		18	Norway	*	*	*	
64	Azerbaijan			*			38	Pakistan			*	*
69	Bangladesh			*	*		51	Peru			*	*
112	Belarus			*	*		608	Philippines				*
7	Belgium	*	*		*		25	Poland		*	*	*
28	Brazil		*	*			41	Portugal		*		
2	Britain	*	*		*		630	Puerto Rico				*
36	Bulgaria		*		*		807	Macedonia				*
12	Canada	*	*		*		61	Moldova			*	*
30	Chile		*	*	*		37	Romania		*	*	
39	China		*	*	*		50	Russia		*	*	*
6	Denmark	*	*		*		15	S Africa		*		*
68	Dominic Rep			*			24	S Korea		*	*	*
818	Egypt				*		703	Slovakia		*		*
222	El Salvador			*			35	Slovenia		*		*
48	Estonia			*	*		8	Spain	*	*	*	*
23	Finland		*	*	*		19	Sweden	*	*	*	*
1	France	*	*		*		26	Switzerland			*	
62	Georgia			*			40	Taiwan			*	
56	Ghana			*			44	Turkey		*	*	*
16	Hungary		*		*		800	Uganda				*
21	Iceland	*			*		49	Ukraine			*	*
32	India		*	*	*		834	Tanzania				*
360	Indonesia				*		54	Uruguay			*	
9	Ireland	*	*		*		11	USA	*	*	*	*
4	Italy	*	*		*		53	Venezuela			*	*
13	Japan	*	*	*	*		704	Viet Nam				*
400	Jordan				*		3	W Germany	*	*	*	*
47	Latvia		*	*	*		716	Zimbabwe				*
46	Lithuania			*	*							

Table 5.2 Summary measures for the dependent variables

Variable	Sample size	Missingness	Min	Max	Mean	SD
Happiness	171264	2896	-4.7958	4.6012	-0.0058	0.9992
Health	171264	3178	-4.6256	4.4059	0.0000	0.9990
Life Satisfaction	171264	4019	-4.4876	4.2929	0.0010	1.0014

Figure 5.1 Repeated cross-sectional multilevel data structure with response nested within individuals nested within waves nested within countries.



*W1: 1981, W2: 1990, W3: 1995-1997 and W4: 1999-2001. $J_{1\text{-}10}$ indicates the residences of Japan as well as for U for United Kingdom and T for Taiwan. SH is the abbreviation of "Self-reported health"; H: Happiness; and LS: Life satisfaction.

5.3 METHODOLOGICAL FRAMEWORK

Multivariate response data are conveniently incorporated into a multilevel model by creating an extra level "below" the original level-one units to define the multivariate structure (Rasbash *et al.,* 2004). Thus different responses are nested within individuals that are in turn nested within waves nested within countries. The basic form of the model (without any predictor variables) is:

$$Y_{1jkl} = \beta_1 X_{0jkl}SH + (v_{1jkl}SH + \mu_{1kl}\,SH + \gamma_{1l}SH)$$

$$Y_{2jkl} = \beta_2 X_{0jkl}H + (v_{2jkl}H + \mu_{2kl}H + \gamma_{2l}H)$$

$$Y_{3jkl} = \beta_3 X_{0jkl}LS + (v_{3jkl}LS + \mu_{3kl}LS + \gamma_{3l}LS)$$

The response is a 'long vector' with all three outcome variables 'stacked' underneath each other so that Y_{1jkl} is the first response (Self-rated health) for individual j in wave k of country l, the variables SH, H and LS are dummy indicator variables that identify the specific response variables. The β's give the estimated average response across all individuals and countries, the γ's give the differential for countries from the global average, the μ's give the wave differential from the country differential while the v's give the individual differential from the wave differential. At each level, all the random effects (for individuals, waves and countries) are assumed to come from a joint normal distribution. Thus, γ's are assumed to come from a multivariate Gaussian distribution which can be summarised by the following variance-covariance matrix:

$$(\gamma_{1l}, \gamma_{2l}, \gamma_{3l}) \sim N \begin{bmatrix} \sigma^2_{\gamma 1} & \sigma_{\gamma 12} & \sigma_{\gamma 13} \\ \sigma_{\gamma 12} & \sigma^2_{\gamma 2} & \sigma_{\gamma 23} \\ \sigma_{\gamma 13} & \sigma_{\gamma 23} & \sigma^2_{\gamma 3} \end{bmatrix}$$

so that the variances on the main diagonal estimate how each response variable varies between countries, while the co-variances will allow the calculation of the correlation between the responses at the country level. Thus we can assess the extent to which countries with good levels of self-rated health are also those with high levels of happiness, and life satisfaction. The variance-covariance matrix at the wave and individual level will allow us to see the extent to which individuals and waves within countries are correlated in the dimensions of self-rated health, etc.

We can then include variables at each level, such as individual income and country inequality, as specified in detail in Chapter Four. This will allow us to see the relationships between income inequality and the responses after controlling for individual income and average country income. Moreover, the variance co-variance matrices at each level will allow us to assess how much between-country and within-country variability has been accounted for by income variables, while the correlations between individuals and countries will now be conditional on the variables that have been included in the fixed part of the model.

Data will be arranged to allow modelling on four levels: the various outcomes (Self-rated health, Life satisfaction and Happiness) at level one; nested in individuals at level two; nested within different waves at level three; nested within countries at level four. The same analysis strategy as in the last chapter will be applied to the model development. For the random part, however, there is no complex between-country variance for individual income of three categories. This is because with only 69 countries there is not enough degree of freedom to permit estimation of this more complex model. In all the models, individual sampling weights are included so as to provide a representative national sample. In practical terms, three sets of individual weights are stacked to form a column the same length as response indicators and then these weights are declared at level 2 that is the level of the individuals. All the models have been fitted using the MLwin Software (Rasbash et al., 2004) using IGLS estimation to get initial models, then a burn in[1] of 500 and then 50000 simulations. The resultant models can be compared with the DIC diagnostic.

5.4 RESULTS

Table 5.3 presents the results of the multivariate multilevel models in the order of their development and complexity. The first model (Model 1) is a null (empty) model with four levels and a multivariate structure. Because each dependent variable has a variance of approximately 1, it is a simple matter to partition the variability to within and between countries. Thus for Health, 19% of the variability is between countries whereas 81% is within countries between waves between people. The variable with the greatest between-country variation is happiness, and the lowest is for self-rated health. It must be stressed, however, that all three variables show significant between-country differences (chi-square analysis revealed that all were significant at $p = 0.05$). Clearly there is a significant national geography to the three variables.

Turning now to the correlation between the three variables, the multivariate multilevel model allows us to estimate the correlation between countries, between waves and between people. Bold numbers in the random parameters section shows correlations between the three responses of each model in Table 5.3. At the country level, the correlations are much more substantial. Happiness, Self-rated health and Life satisfaction are highly correlated at more than 0.5; so countries with happy people

[1] Typically, 'burn in' samples are initial samples which are not completely valid. This is because the Markov Chain has not converged. So the 'burn in' samples allow you to discard these initial samples. Odd that this is here when used MCMC in last chapter

tend to be places with good self-rated health and high life satisfaction. In Figure 5.2, a matrix plot of country residuals for the three variables shows such a pattern, for example, Puerto Rico (code 630) is in the top ten highest countries on all three responses.

Figure 5.2 The matrix plot of country level residuals for Happiness, Self-rated health and Life satisfaction variables

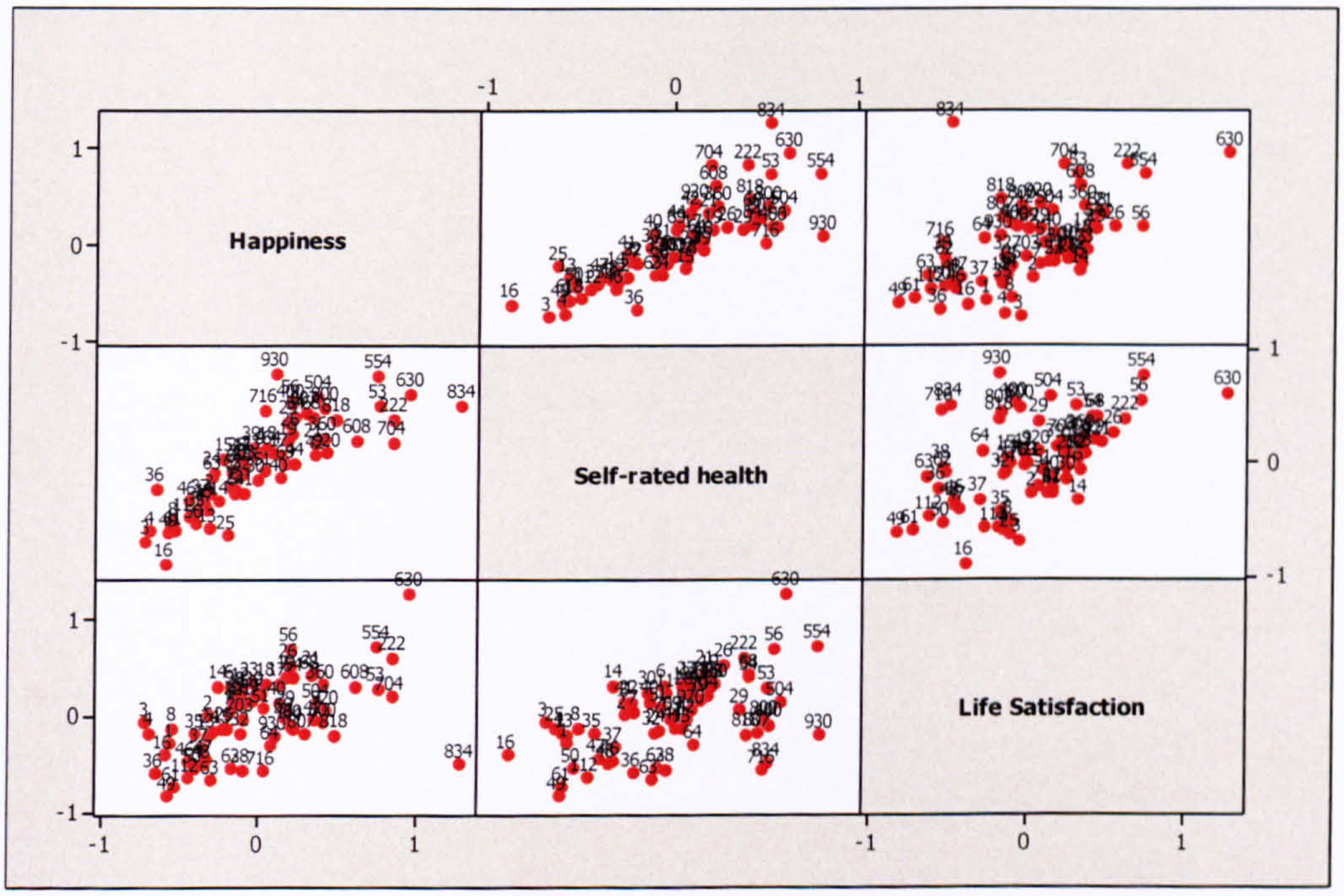


Table 5.4 lists in detail the country residuals of the top ten and bottom ten countries from the global average. Countries such as Puerto Rico and New Zealand are within the top ten countries that are high on self-rated health, life satisfaction and happiness. However, Tanzania has the high levels of happiness and health but low levels of life satisfaction. Moreover, countries from the former Soviet Bloc such as Ukraine, Hungary and the Russian Federation, have the lowest happiness, health and life satisfaction. In addition, some OECD high economic countries surprisingly reported themselves unhappy and not very healthy in their countries such as Spain, Italy, France and West Germany.

Blank Page

Table 5.3 Results of fixed and random part of the Multivariate analytical models and Deviance information criterion (MCMC)

Fixed Parameters	Model 1 happy (S.E.)	Model 1 health (S.E.)	Model 1 life (S.E.)	Model 2 happy (S.E.)	Model 2 health (S.E.)	Model 2 life (S.E.)	Model 3 happy (S.E.)	Model 3 health (S.E.)	Model 3 life (S.E.)
Constant	-0.090 (0.064)*	-0.026 (0.041) *	0.053 (0.051) *	0.128 (0.058)	0.667 (0.049)	0.084 (0.048)*	0.038 (0.024)*	0.593 (0.024)	0.474 (0.029)
Wave 1 (base)									
Wave 2	0.037 (0.057)*[2]	0.034 (0.030)*	0.029 (0.048)*	0.062 (0.046)*	0.101 (0.030)	0.070 (0.049)*	0.052 (0.040)*	0.061 (0.029)	-0.101 (0.029)
Wave 3	0.197 (0.063)	0.091 (0.034)	-0.059 (0.049) *	0.236 (0.056)	0.176 (0.035)	-0.023 (0.050)*	0.233 (0.046)	0.140 (0.030)	-0.232 (0.029)
Wave 4	0.292 (0.071)	0.203 (0.041)	-0.076 (0.055)*	0.302 (0.060)	0.277 (0.039)	-0.052 (0.064)*	0.303 (0.046)	0.251 (0.038)	-0.292 (0.039)
Individual predictors									
Female				0.111 (0.012)	-0.074 (0.012)	0.087 (0.012)	0.110 (0.012)	-0.073 (0.010)	0.089 (0.011)
Age (centred around 40)				-0.002 (0.000)	-0.015 (0.000)	0.001 (0.000)	-0.002 (0.000)	-0.015 (0.000)	0.001 (0.000)
Female*Age				-0.001 (0.000)	-0.001 (0.000)	-0.001 (0.000)	-0.001 (0.000)	-0.001 (0.000)	-0.001 (0.000)
Married status									
(Couple as the base)									
Widowed/ Separate/Divorced				-0.316 (0.008)	-0.087 (0.007)	-0.189 (0.008)	-0.316 (0.008)	-0.087 (0.007)	-0.189 (0.008)
Single				-0.157 (0.006)	-0.016 (0.006)	-0.070 (0.006)	-0.156 (0.006)	-0.015 (0.006)	-0.069 (0.006)
Unknown				-0.127 (0.051)	-0.020 (0.048) *	-0.094 (0.053)*	-0.126 (0.050)	-0.018 (0.050)*	-0.094 (0.053)*
Income (5&6 as the base)									
1&2				-0.229 (0.007)	-0.235 (0.007)	-0.287 (0.008)	-0.230 (0.007)	-0.235 (0.007)	-0.287 (0.008)
3&4				-0.090 (0.007)	-0.096 (0.007)	-0.128 (0.007)	-0.090 (0.007)	-0.095 (0.007)	-0.128 (0.007)
7&8				0.060 (0.007)	0.078 (0.007)	0.087 (0.008)	0.060 (0.007)	0.079 (0.007)	0.087 (0.008)
9&10				0.102 (0.009)	0.156 (0.009)	0.155 (0.010)	0.102 (0.009)	0.156 (0.009)	0.155 (0.010)
Unknown				-0.039 (0.008)	-0.050 (0.008)	-0.008 (0.009) *	-0.040 (0.009)	-0.050 (0.008)	-0.009 (0.009)*
Country predictors									
GDP (centred around 10K)							-0.038 (0.006)	-0.034 (0.004)	0.040 (0.003)
Wilkinson							0.025 (0.051)*	0.053 (0.040)*	-0.229 (0.038)
(below 5K as the base)									
GDP*Wilkinson							0.040 (0.007)	0.038 (0.006)	-0.017 (0.004)

2 All estimates are significant at 0.05 probability level, except those marked by *, which have a probability greater than 0.05.

EHII (centred around 40)									
EHII*Wilkinson									
Random parameters									
Level 4: Between countries	**Happiness**	**Self-rated health**	**Life satisfaction**	**Happiness**	**Self-rated health**	**Life satisfaction**	**Happiness**	**Self-rated health**	**Life satisfaction**
Happiness	0.193 (0.038)			0.200 (0.040)			0.199 (0.041)		
	1.000[3]			**1.000**			**1.000**		
Self-rated health	0.141 (0.030)	0.162 (0.030)		0.136 (0.029)	0.143 (0.027)		0.132 (0.030)	0.144 (0.029)	
	0.798	**1.000**		**0.806**	**1.000**		**0.788**	**1.000**	
Life satisfaction	0.099 (0.028)	0.082 (0.024)	0.171 (0.032)	0.104 (0.029)	0.083 (0.023)	0.164 (0.032)	0.126 (0.032)	0.097 (0.025)	0.144 (0.031)
	0.545	**0.493**	**1.000**	**0.576**	**0.538**	**1.000**	**0.741**	**0.676**	**1.000**
Level 3: Between waves							Happiness	Self-rated health	Life satisfaction
Happiness	0.023 (0.005)			0.024 (0.005)			0.024 (0.005)		
	1.000			**1.000**			**1.000**		
Self-rated health	0.005 (0.003)	0.007 (0.002)		0.007 (0.003)	0.008 (0.002)		0.007 (0.003)	0.009 (0.002)	
	0.411	**1.000**		**0.507**	**1.000**		**0.477**	**1.000**	
Life satisfaction	0.009 (0.003)	0.001 (0.002)*	0.015 (0.003)	0.0013 (0.005)	0.004 (0.002)	0.021 (0.006)	0.012 (0.004)	0.003 (0.002)*	0.020 (0.006)
	0.494	**0.108**	**1.000**	**0.563**	**0.292**	**1.000**	**0.564**	**0.227**	**1.000**
Level 2: Between individuals									
Happiness	0.809 (0.003)			0.782 (0.003)			0.782 (0.003)		
	1.000			**1.000**			**1.000**		
Self-rated health	0.247 (0.002)	0.844 (0.003)		0.217 (0.002)	0.759 (0.003)		0.217 (0.002)	0.759 (0.003)	
	0.299	**1.000**		**0.282**	**1.000**		**0.282**	**1.000**	
Life satisfaction	0.329 (0.002)	0.196 (0.002)	0.874 (0.003)	0.307 (0.002)	0.177 (0.002)	0.852 (0.003)	0.307 (0.002)	0.177 (0.002)	0.852 (0.003)
	0.392	**0.228**	**1.000**	**0.377**	**0.221**	**1.000**	**0.377**	**0.221**	**1.000**
Deviance information criterion (DIC, 69 countries)			133286.35			1308616.07			1308611.04
Change in DIC			--			2470.28			5.03

3 Numbers marked in bold are correlation of between countries (level 4), within country between waves (level 3) and within wave between individuals (level 2) for these three responses of happiness, health and life satisfaction.

Table 5.3 Results of fixed and random part of the Multivariate analytical models and Deviance information criterion (MCMC) – continued

Fixed Parameters	Model 4 ^{happy} (S.E.)	Model 4 ^{health} (S.E.)	Model 4 ^{life} (S.E.)	Model 5 ^{happy} (S.E.)	Model 5 ^{health} (S.E.)	Model 5 ^{life} (S.E.)
Constant	0.015 (0.174)*	0.073 (0.144)	0.538 (0.183)	0.087 (0.210)*	-0.102 (0.189)	0.453 (0.237)
Wave 1 (base)						
Wave 2	0.044 (0.070)*	0.103 (0.056)*	-0.142 (0.074)	0.053 (0.086)*	-0.103 (0.066)*	-0.133 (0.084)
Wave 3	0.187 (0.098)*	0.179 (0.081)*	-0.323 (0.106)	0.203 (0.116)*	0.177 (0.094)*	-0.303 (0.120)
Wave 4	0.258 (0.096)	0.301 (0.092)	-0.375 (0.115)	0.289 (0.129)	0.306 (0.107)	-0.345 (0.134)
Individual predictors						
Female	0.060 (0.005)	-0.104 (0.004)	0.045 (0.005)	0.060 (0.005)	-0.104 (0.005)	0.044 (0.005)
Age (centred around 40)	-0.002 (0.000)	-0.015 (0.000)	0.001 (0.000)	-0.002 (0.000)	-0.015 (0.000)	0.001 (0.000)
Female*Age	-0.001 (0.000)	0.001 (0.000)	-0.001 (0.000)	-0.001 (0.000)	-0.001 (0.000)	-0.001 (0.000)
Married status (Couple as the base)						
Widowed/ Separate/Divorced	-0.316 (0.008)	-0.091 (0.006)	-0.200 (0.008)	-0.320 (0.008)	-0.092 (0.008)	-0.205 (0.008)
Single	-0.161 (0.006)	-0.019 (0.006)	-0.077 (0.006)	-0.167 (0.006)	-0.020 (0.006)	-0.083 (0.006)
Unknown	-0.133 (0.051)	0.020 (0.050)*	-0.097 (0.053)*	-0.135 (0.052)	0.020 (0.049)*	-0.100 (0.053)*
Income (5&6 as the base)						
1&2	-0.229 (0.008)	-0.237 (0.008)	-0.287 (0.008)	-0.226 (0.008)	-0.237 (0.008)	-0.286 (0.008)
3&4	-0.087 (0.007)	-0.096 (0.007)	-0.127 (0.007)	-0.088 (0.007)	-0.096 (0.007)	-0.127 (0.007)
7&8	0.057 (0.008)	0.080 (0.008)	0.072 (0.008)	0.060 (0.008)	0.081 (0.008)	0.076 (0.008)
9&10	0.099 (0.010)	0.162 (0.010)	0.140 (0.010)	0.114 (0.010)	0.168 (0.010)	0.159 (0.011)
Unknown	-0.029 (0.008)	-0.046 (0.008)	-0.007 (0.009)*	-0.031 (0.009)	-0.049 (0.008)	-0.005 (0.009)*
Country predictors						
GDP (centred around 10K)	-0.007 (0.021)*	0.024 (0.016)*	0.041 (0.019)	-0.019 (0.024)*	-0.027 (0.020)*	0.029 (0.025)*
Wilkinson (below as the base)	-0.052 (0.141)*	0.091 (0.117)*	-0.214 (0.151)*	-0.090 (0.181)*	0.164 (0.174)*	-0.089 (0.188)*
GDP*Wilkinson	0.006 (0.021)*	0.021 (0.015)*	-0.018 (0.019)*	0.018 (0.022)*	0.024 (0.018)*	-0.007 (0.023)*
EHII (centred around 40)	0.022 (0.011)	0.013 (0.009)*	0.023 (0.011)	0.023 (0.010)	0.015 (0.008)*	0.023 (0.009)
EHII*Wilkinson	-0.007 (0.013)*	-0.004 (0.010)*	-0.009 (0.013)*	-0.008 (0.013)*	-0.006 (0.010)*	-0.011 (0.014)*
Micro-Macro interaction terms						
EHII*1&2				-0.008 (0.001)	-0.001 (0.001)*	-0.007 (0.001)
EHII*3&4				-0.002 (0.01)	-0.001 (0.001)*	-0.003 (0.001)
EHII*7&8				0.002 (0.001)	0.001 (0.001)*	-0.004 (0.001)
EHII*9&10				0.008 (0.002)	0.004 (0.001)	0.011 (0.002)
EHII*Unknown				-0.002 (0.001)	-0.002 (0.001)	0.002 (0.001)

Random parameters						
Level 4: Between countries	**Happiness**	**Self-rated health**	**Life satisfaction**	**Happiness**	**Self-rated health**	**Life satisfaction**
Happiness	0.148 (0.033)			0.152 (0.033)		
	1.000			**1.000**		
Self-rated health	0.114 (0.027)	0.135 (0.027)		0.115 (0.026)	0.136 (0.027)	
	0.807	**1.000**		**0.802**	**1.000**	
Life satisfaction	0.114 (0.028)	0.087 (0.023)	0.135 (0.029)	0.121 (0.029)	0.090 (0.024)	0.143 (0.031)
	0.810	**0.645**	**1.000**	**0.820**	**0.642**	**1.000**
Level 3: Between waves	Happiness	Self-rated health	Life satisfaction	Happiness	Self-rated health	Life satisfaction
Happiness	0.023 (0.005)			0.022 (0.005)		
	1.000			**1.000**		
Self-rated health	0.005 (0.003)	0.008 (0.002)		0.005 (0.002)	0.008 (0.002)	
	0.414	**1.000**		**0.418**	**1.000**	
Life satisfaction	0.010 (0.004)	0.001 (0.002)*	0.016 (0.004)	0.009 (0.003)	0.001 (0.002)*	0.014 (0.004)
	0.525	**0.133**	**1.000**	**0.494**	**0.100**	**1.000**
Level 2: Between individuals						
Happiness	0.774 (0.003)			0.773 (0.003)		
	1.000			**1.000**		
Self-rated health	0.215 (0.002)	0.754 (0.003)		0.214 (0.002)	0.754 (0.003)	
	0.281	**1.000**		**0.281**	**1.000**	
Life satisfaction	0.309 (0.002)	0.177 (0.002)	0.837 (0.003)	0.308 (0.002)	0.177 (0.002)	0.836 (0.003)
	0.383	**0.223**	**1.000**	**0.383**	**0.223**	**1.000**
Deviance information criterion			1174325.71			1174150.52
(DIC, 61 countries)						
Change in DIC			134285.33			175.19

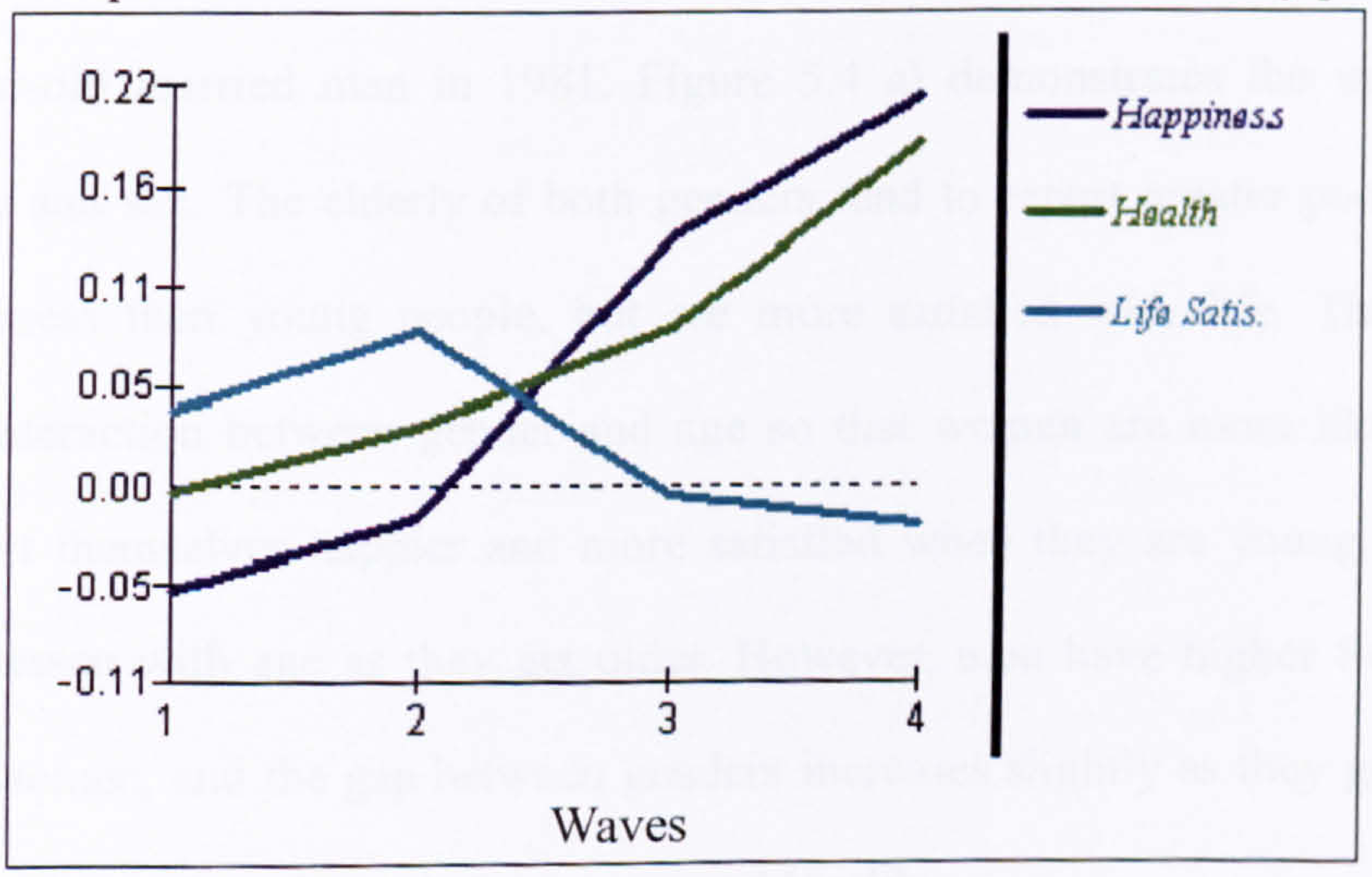
Table 5.4 Country differentials from global averages in Happiness, Self-rated health and Life satisfaction; ten best and ten worst countries for each variable

Rank	Country	Happiness	Country	Self-rated Health	Country	Life satisfaction
1	Tanzania	1.582	Albania	0.946	Puerto Rico	1.356
2	Puerto Rico	1.174	New Zealand	0.814	New Zealand	0.737
3	VietNam	1.117	Puerto Rico	0.729	Ghana	0.706
4	El Salvador	1.006	Morocco	0.722	El Salvador	0.666
5	Venezuela	0.868	Jordan	0.660	Switzerland	0.550
6	Philippines	0.837	Tanzania	0.642	Iceland	0.536
7	New Zealand	0.816	Uganda	0.630	Sweden	0.454
8	Egypt	0.639	Zimbabwe	0.622	Dominic Rep	0.423
9	Algeria	0.622	Macedonia	0.589	Uruguay	0.398
10	Indonesia	0.569	Ghana	0.583	Finland	0.391
…	…	…	…	…	…	…
60	Britain	-0.481	Moldova	-0.550	Tanzania	-0.531
61	Romania	-0.515	Russia	-0.572	Estonia	-0.538
62	Slovenia	-0.545	Spain	-0.580	Bulgaria	-0.579
63	Ukraine	-0.574	Japan	-0.587	Georgia	-0.600
64	Spain	-0.593	Ukraine	-0.629	Pakistan	-0.650
65	Hungary	-0.705	France	-0.666	Belarus	-0.682
66	France	-0.727	Poland	-0.671	Zimbabwe	-0.691
67	W Germany	-0.827	Italy	-0.677	Armenia	-0.732
68	Italy	-0.850	W Germany	-0.754	Moldova	-0.786
69	Bulgaria	-0.892	Hungary	-0.977	Ukraine	-0.899

There is relatively small within country between-waves variations which mean that the countries are quite similar between different waves. At this level, the correlation between Self-rated health and Happiness is still high, but the correlation is very low between Self-rated health and Life satisfaction, even lower than the correlation between individuals. So that changes over time in Self-rated health is not correlated with changes over time in Life satisfaction. It is not surprising that there are very large within-wave between-individuals variations. The correlations between the three variables are much smaller. In other words, knowing about people who report themselves in a good health does not tell you in a great deal about whether they are happy or satisfied with their life. But all the correlations are still trivially significant ($P < 0.000$) due to the large sample size.

The time variables of different waves are needed in the fixed part of the model to describe the general trend, as the z-ratios for the fixed part coefficients of all the waves in each of the dependents exceeds the 0.05 cut-off of 2. Figure 5.3 plots the predicted trend of each outcome variable across the study period. In general, Happiness and Self-rated health has similar overall increasing patterns, while Life satisfaction has overall downward trends. However, we should not make too much of this because of the issue of different countries in the sample at different waves. There is also danger here of results being significant but based on very large sample size, but not being substantively important.

Figure 5.3 The predicted trend of each outcome variable across the study period



Model 2 examines the relationships between the three dependent variables and the individual variables: age, sex, marital status and income quintile. The change in the DIC between model 1 and model 2 shows the importance of these variables in accounting for variations in the response variables. Table 5.5 summaries all the joint chi-square testing of each of the continuous predictors or each set of categorical dependent variable; e.g. for the sum of five income groups for happiness is 1830.03. All the effects are significant. Married status has the most significant influence on Happiness, as does individual income on Life satisfaction, and age for Self-rated health.

Table 5.5 The joint chi-square for the testing of each set of individual variables for each dependent variable

	Happiness	Health	Life Satisfaction
Age	118.72	4738.81	35.80
Sex	88.34	39.11	50.06
Income groups	1830.03	2460.56	3021.42
Married status	2287.50	140.13	663.88

A series of graphs showing the detail of individual effects separately for each outcome is given in Figure 5.4. These are shown in relation to the base category that is middle class 40-year-old married man in 1981. Figure 5.4 a) demonstrates the individual effect of age and sex. The elderly of both genders tend to report greater poor health and unhappiness than young people, but are more satisfied with life. There is a significant interaction between gender and age so that women are more likely than men to report themselves happier and more satisfied when they are young, but the differences lessen with age as they get older. However, men have higher Self-rated health than women, and the gap between genders increases slightly as they get older. The biggest effect is the relation between age and Health.

Figure 5.4 b), shows the impact for marital status. Respondents in the marital situation of separated, widowed and divorced (SWD), are more likely to report a worse situation than couples and singles; this is especially the case for Happiness and then

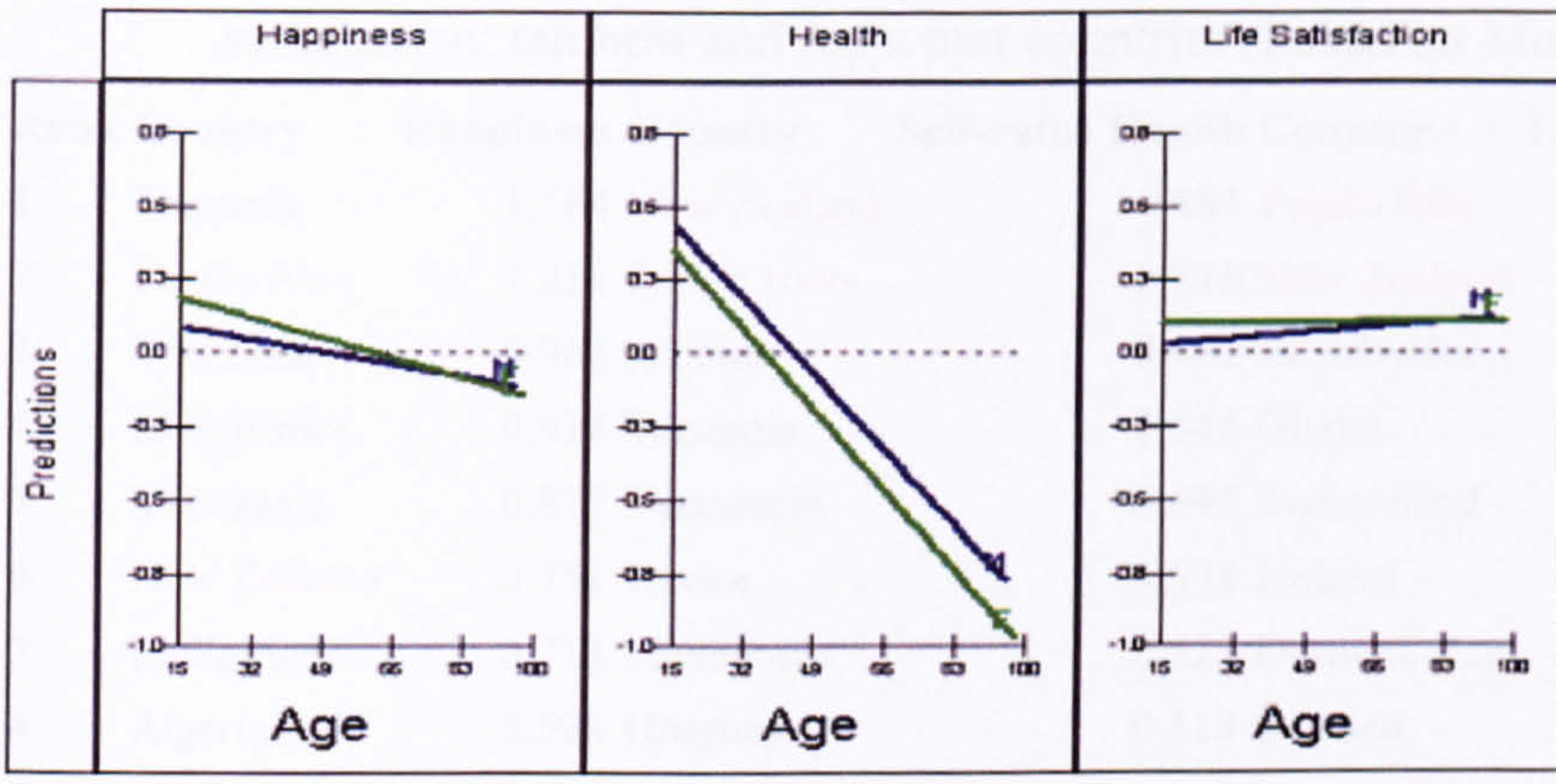
Life satisfaction. Single people experience less happiness, poorer health and less life satisfaction than couples, but the worst results on all three responses are for the SWD group. This result, confirmed by Veenhoven (1989), in summary is that singles are typically less happy than the married, and that the widowed and the divorced are particularly inclined to unhappiness. This difference has always been taken to show that the presence of a partner makes life more satisfying; in other words that marriage brings happiness. Moreover, where marital status is Unknown, there is no difference in terms of health compared with couples; the unknown group is as happy as single people, but is less satisfied with their life than couples or singles.

There is a very consistent 'dose-response' relationship between each of the three outcomes and individual income. Individual income affects them all to a similar degree and in a similar way. As income goes from the wealthiest quintile (9&10) to the lowest quintile (1&2), the less he/she tends to be happy, healthy, and satisfied with himself/herself in life, with the largest effects being found for Life satisfaction (see Figure 5.4 c).

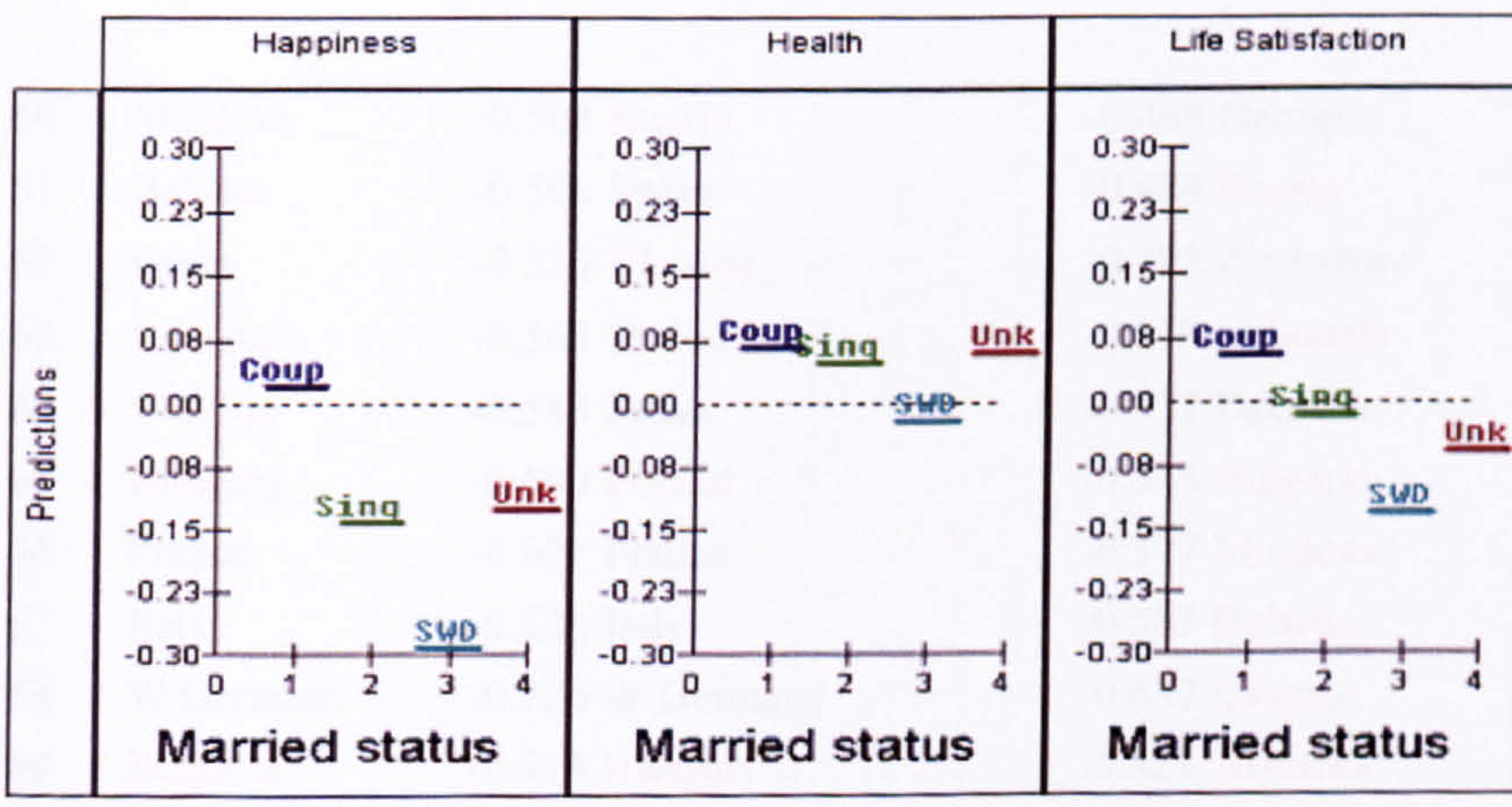
In the random part of the model, the correlation between the dependent variables in general has increased in the country level, and decreased in the individual level (see Model 2 in Table 5.3), but all these changes are relatively small and insubstantial. There remain considerable between-country variations that do not attenuate substantially when individual characteristics are included. In other words, these individual characteristics are not significantly different from country to country. Table 5.6 shows that the country differences for the top ten and bottom ten countries are similar to the previous model with Puerto Rico and New Zealand still within the best ten countries of all outcome variables, and the ten worst countries still dominated by the Eastern European block and some OECD countries.

Figure 5.4 individual effects

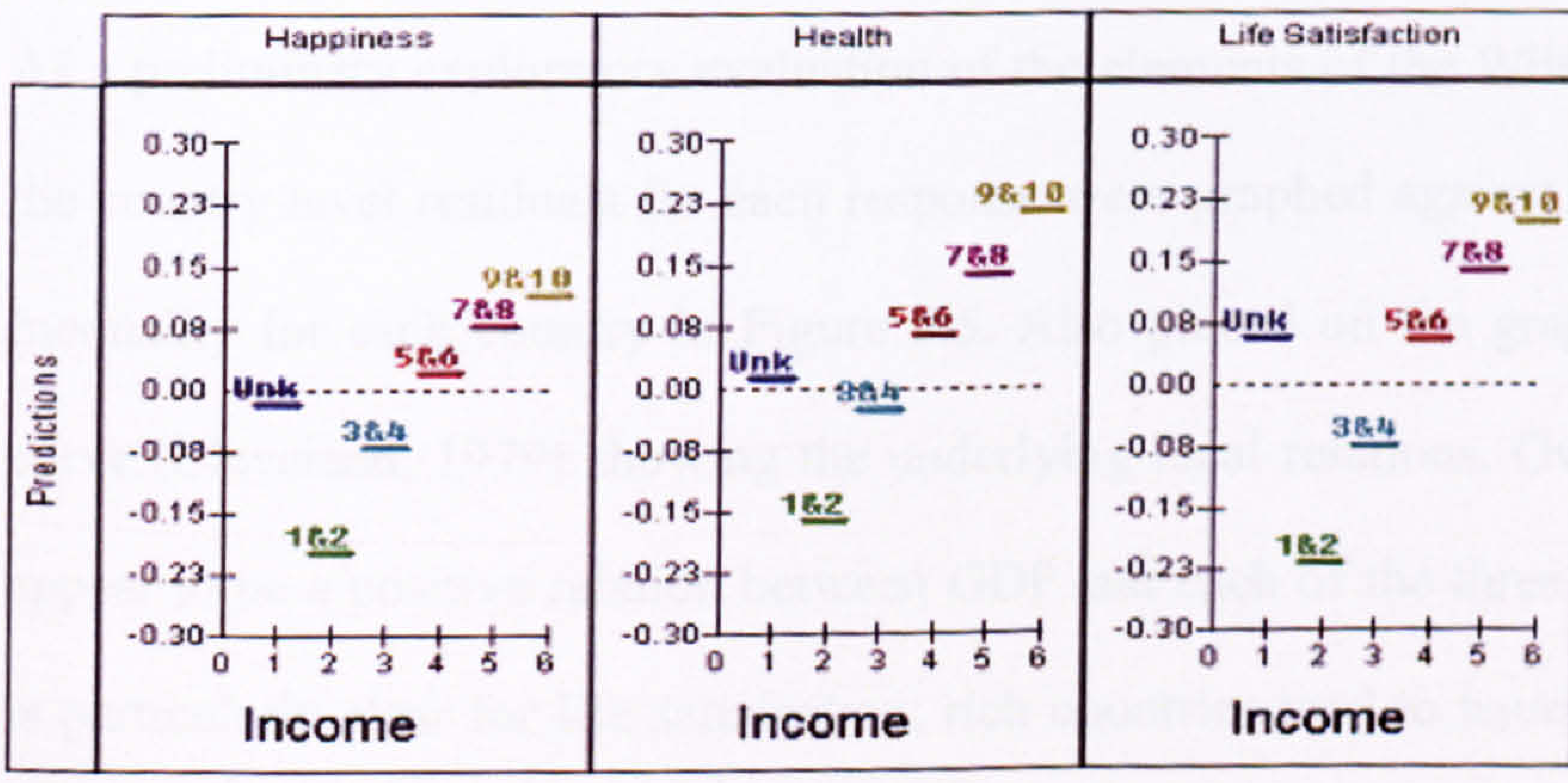
a) Age and Sex



b) Marital status



c) Income groups



Furthermore, there is little change in their ranking in Table 5.6. Zimbabwe has been found to be a country experiencing very good health but lower Life satisfaction.

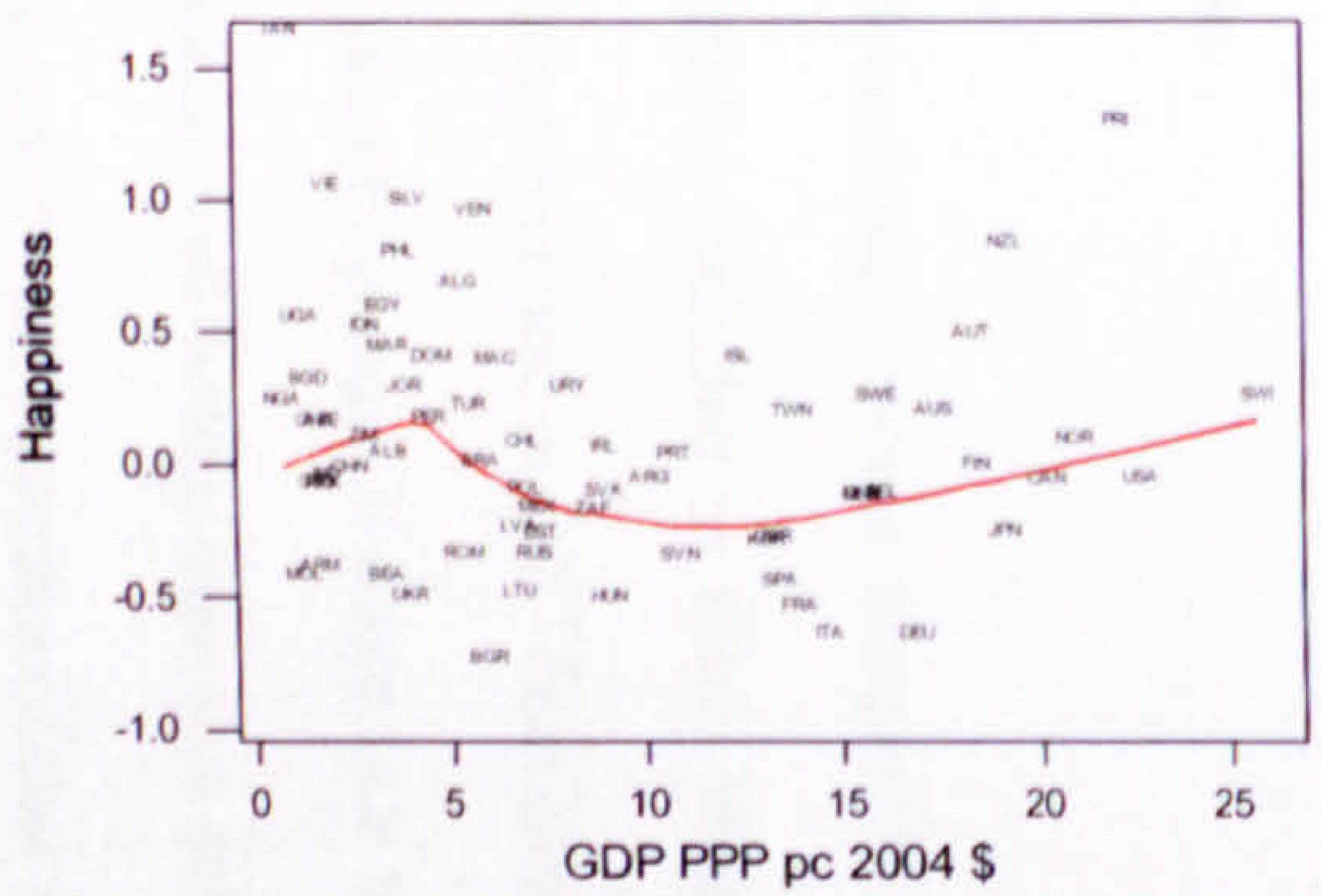
Table 5.6Country differentials from the global average in happiness, health and life satisfaction; ten best and ten worst countries (based on Model 2)

Rank	Country	Happiness	Country	Self-rated Health	Country	Life Satisfaction
1	Tanzania	1.564	New Zealand	0.881	Puerto Rico	1.512
2	Puerto Rico	1.211	Puerto Rico	0.816	New Zealand	0.692
3	Viet Nam	0.968	Albania	0.801	El Salvador	0.629
4	El Salvador	0.914	Tanzania	0.546	Ghana	0.628
5	Venezuela	0.877	Venezuela	0.545	Switzerland	0.519
6	New Zealand	0.751	Jordan	0.534	Iceland	0.490
7	Philippines	0.718	Macedonia	0.523	Dominic Rep	0.448
8	Algeria	0.598	Uruguay	0.518	Sweden	0.428
9	Egypt	0.511	Morocco	0.487	Venezuela	0.409
10	Uganda	0.465	Zimbabwe	0.486	Uruguay	0.394
60	Moldova	-0.502	Russia	-0.468	Georgia	-0.511
61	Belarus	-0.506	Spain	-0.484	Russia	-0.561
62	Spain	-0.525	Ukraine	-0.493	Zimbabwe	-0.568
63	Lithuania	-0.568	Moldova	-0.518	Lithuania	-0.595
64	Ukraine	-0.583	Japan	-0.537	Pakistan	-0.597
65	Hungary	-0.590	Poland	-0.545	Bulgaria	-0.674
66	France	-0.621	France	-0.577	Moldova	-0.707
67	Italy	-0.723	Italy	-0.583	Belarus	-0.712
68	W Germany	-0.726	W Germany	-0.637	Ukraine	-0.857
69	Bulgaria	-0.814	Hungary	-0.821	Armenia	-0.858

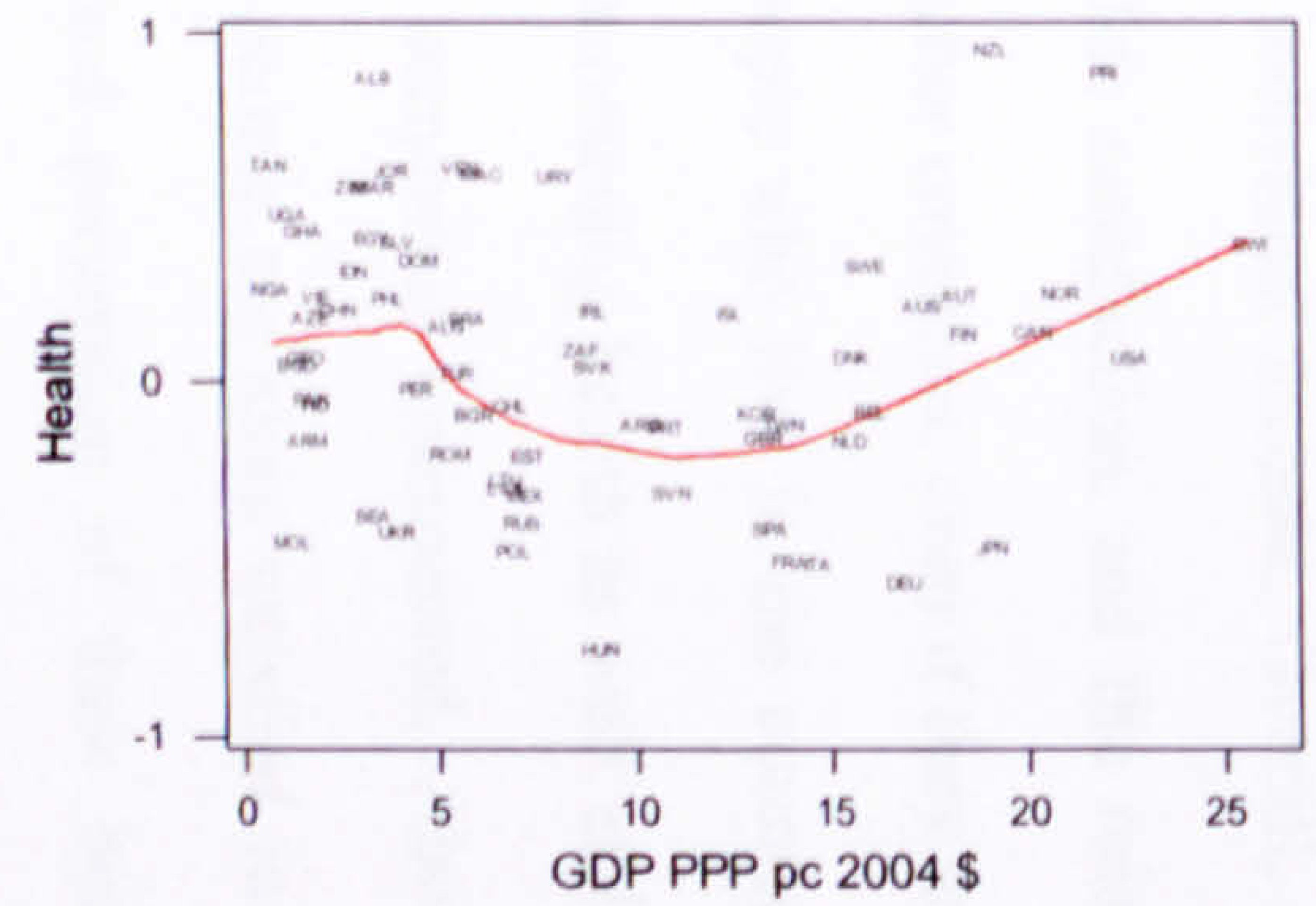
As a preliminary exploratory evaluation of the elements of the Wilkinson hypothesis, the country level residuals for each response were graphed against average GDP and Inequality for each country in Figure 5.5. Also placed on the graph is a LOWESS curve (Cleveland, 1979) showing the underlying local relations. Overall there would appear to be a positive relation between GDP and each of the three responses, which is particularly clear for life satisfaction; rich countries tend to have high average life satisfaction.

Figure 5.5 Underlying trends of between country GDP PPP pc and estimated household income inequality (EHII), respectively, and country residuals of each outcome variable

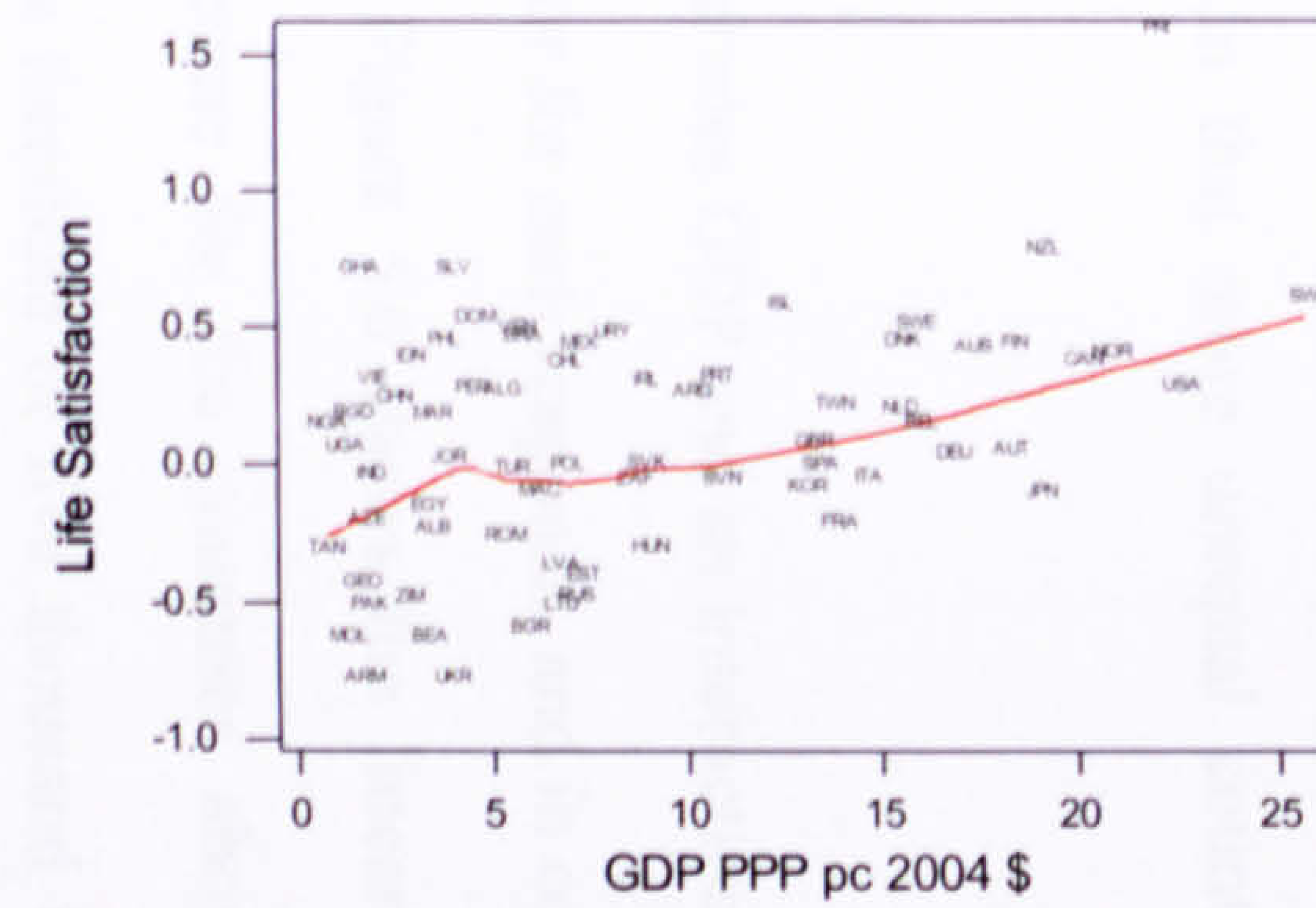
a) Happiness



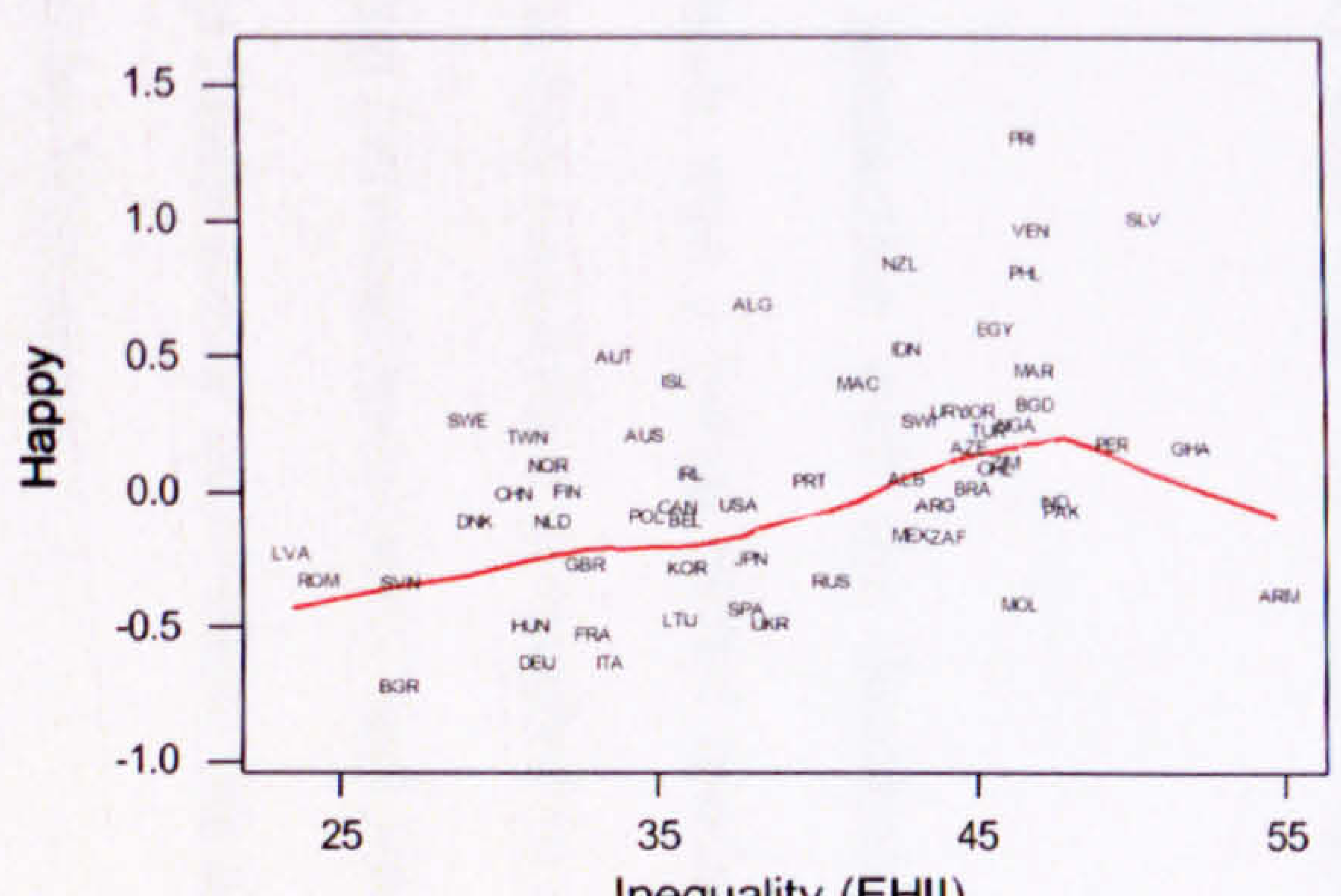
b) Self-rated health



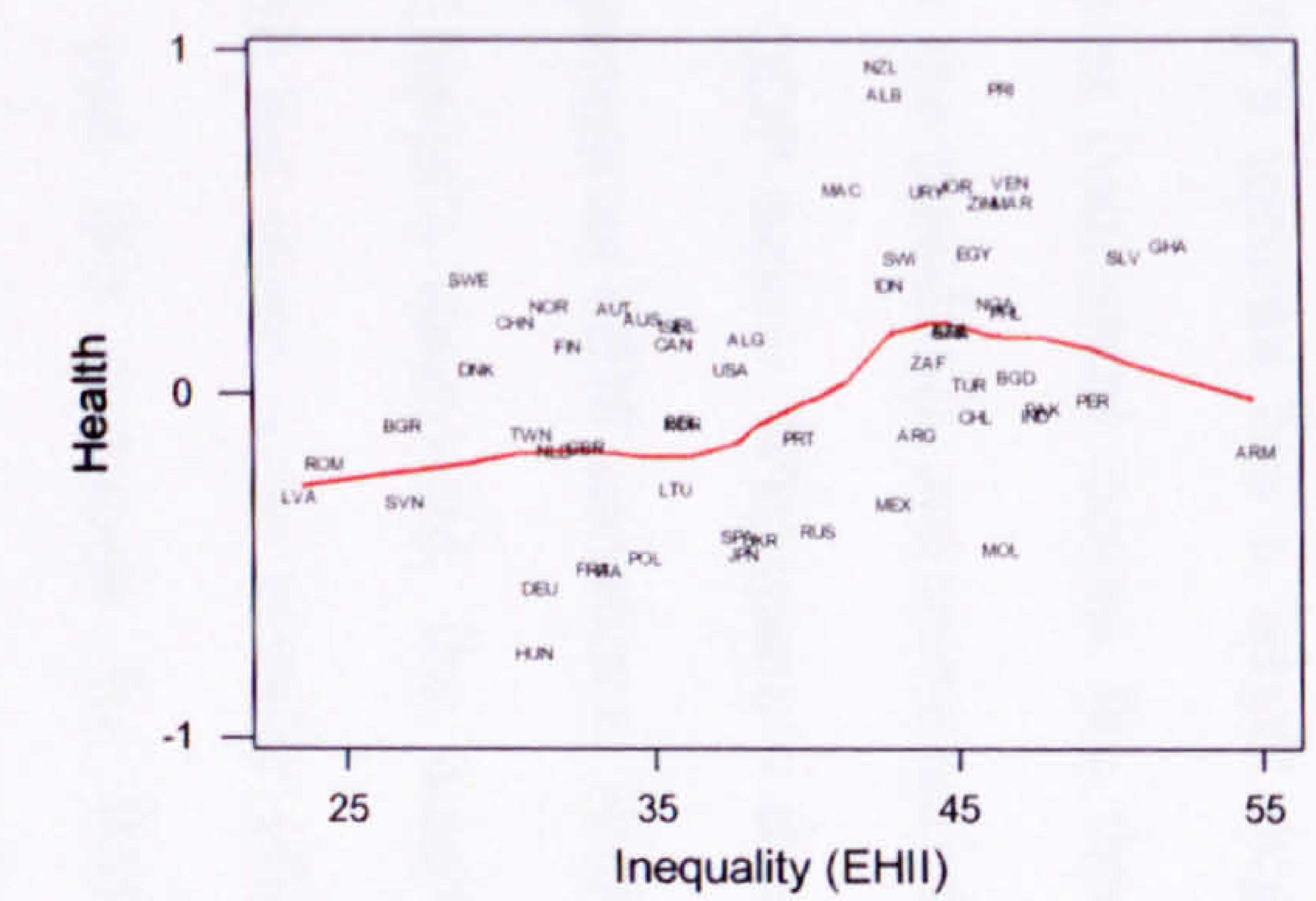
c) Life satisfaction



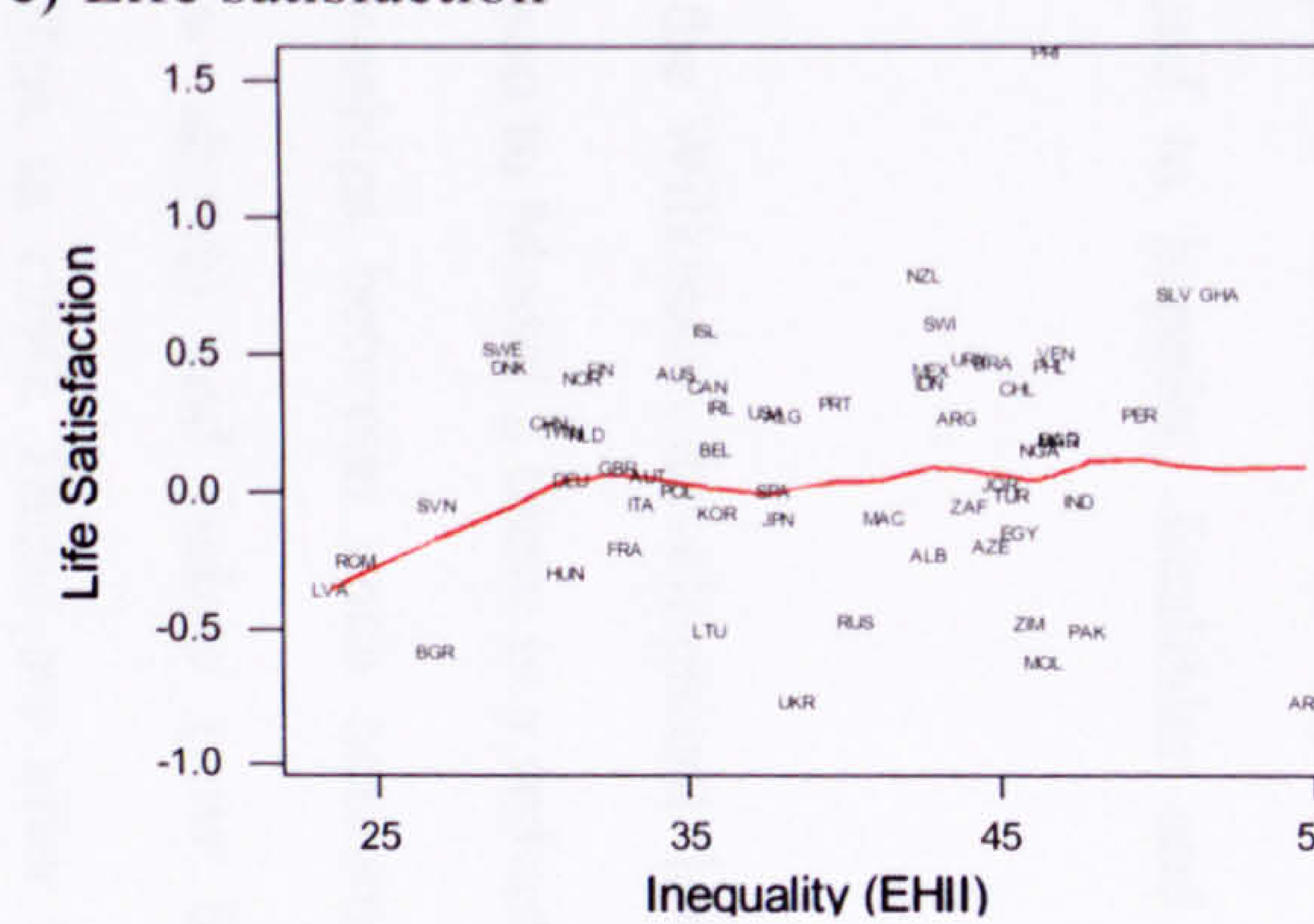
a) Happiness



b) Self-rated health



c) Life satisfaction



But once again the overall results for inequality contradict to the Wilkinson hypothesis in that more unequal countries tend to happier, healthier and more satisfied.

Model 3 includes GDP and an interaction with the Wilkinson development threshold as a predictor for each response, and in comparison to Model 2 there is a reduction in DIC of 5. Figure 5.6 shows the linear relationships between each outcome and country income for the countries above (row above) and below (row below) Wilkinson's threshold of five thousand US dollars in 1990. These are after taking account of individual income. Overall, a country's income has a larger degree of impact on life satisfaction, and the relationship is consistent across the threshold. People in richer countries, even if they are below the threshold, are more satisfied. In contrast, people who live in the counties with GDP below Wilkinson's threshold, experience worsening health as well as less happiness as GDP increases. Both these results are somewhat unexpected. Above Wilkinson's threshold, the wealthier a country is, people are more satisfied in their life but there is no similar effect for happiness and perception of their health. In part this supports the Wilkinson hypothesis; a richer advanced economy country does not result in a happier or healthier one.

Estimated household income inequality and its interaction with the threshold were then additionally included to form Model 4, which was a better fitted model with substantial improvement in DIC than Model 3. The key results are presented in Figure 5.7 and are unambiguous. If there is a relationship between inequality and the three aspects of social well-being, then this relationship (both above and below the income threshold) is exactly the opposite of that suggested by Wilkinson.

Figure 5.6 Separated outcome predictions associated with country income

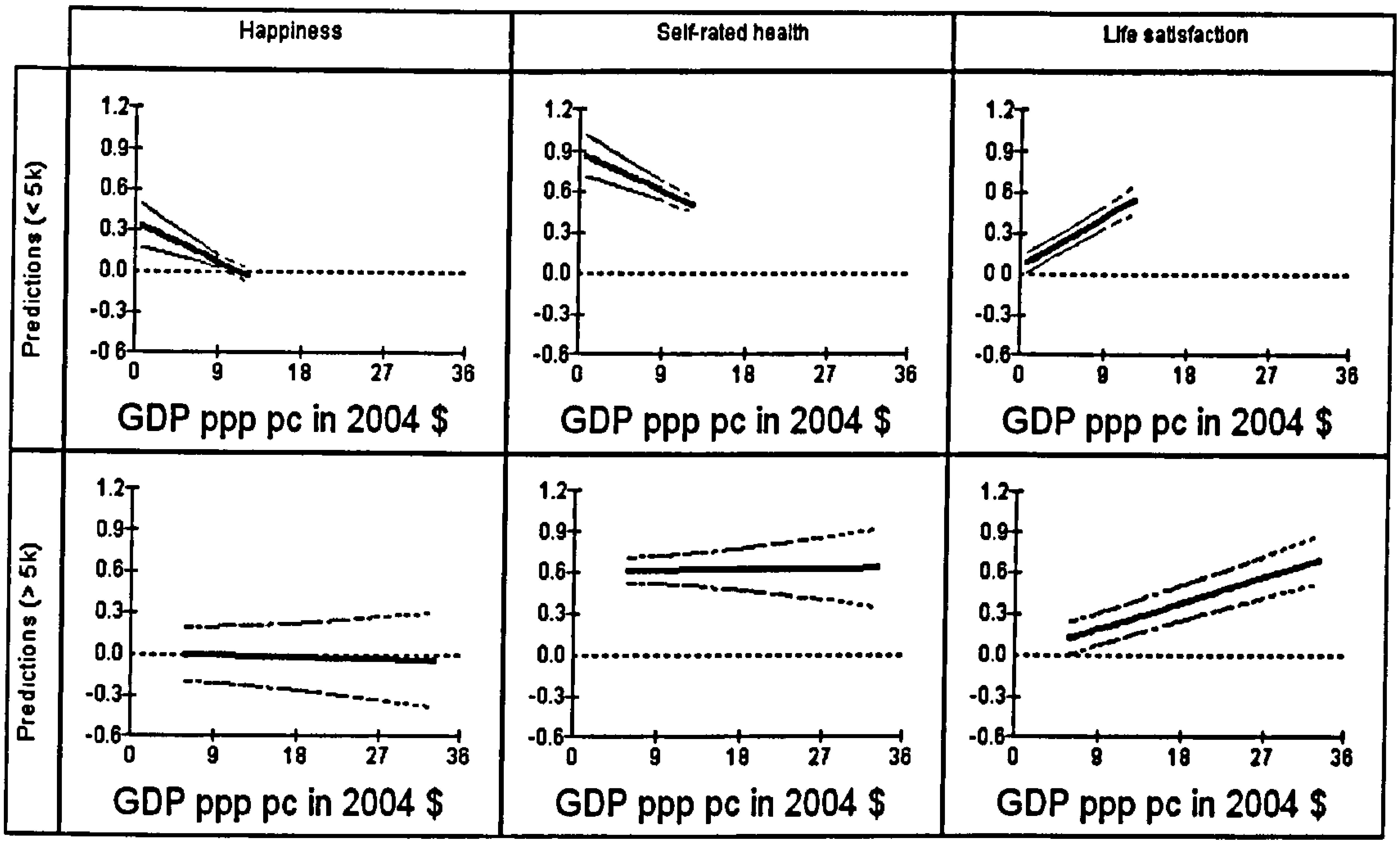
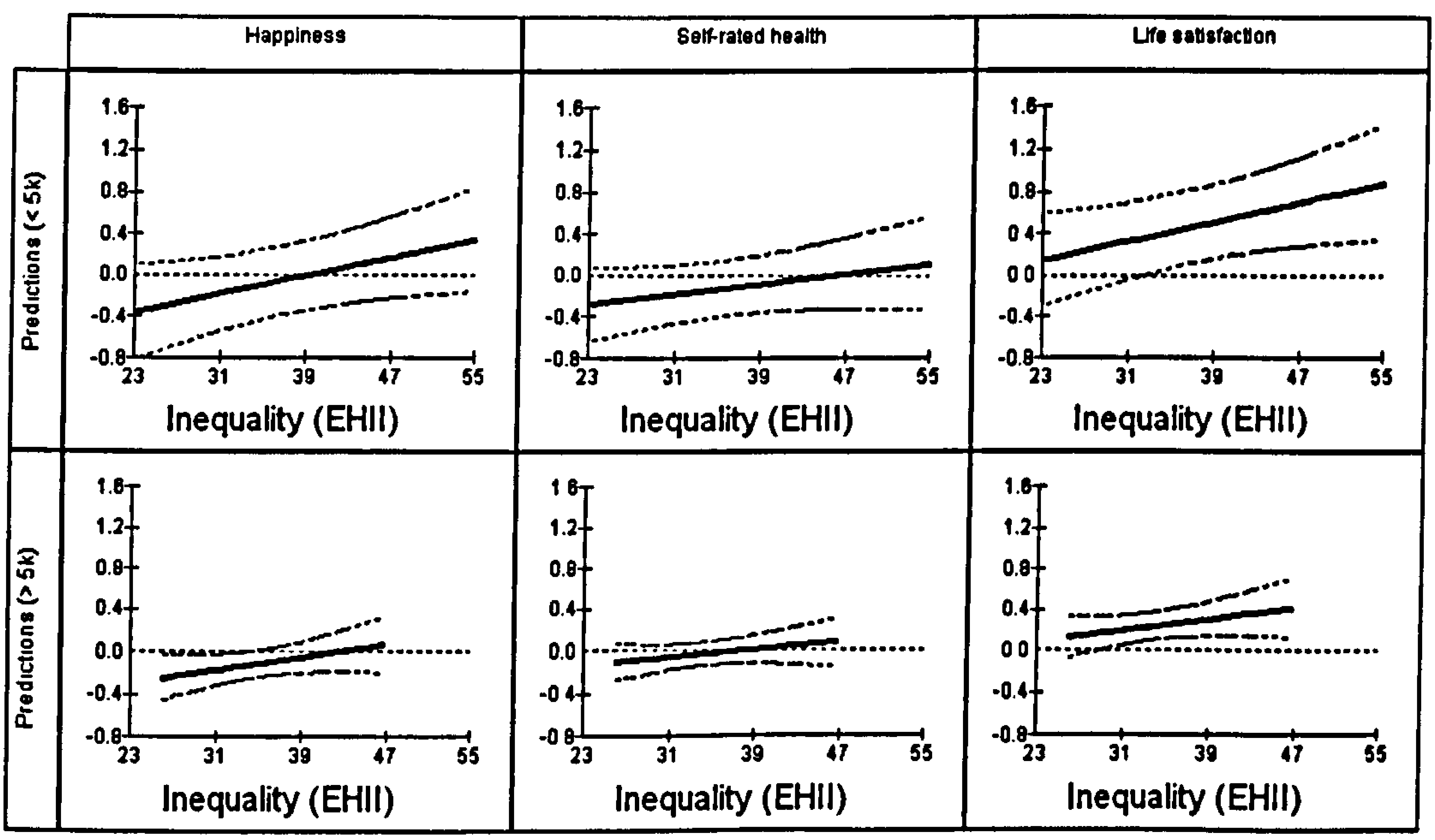


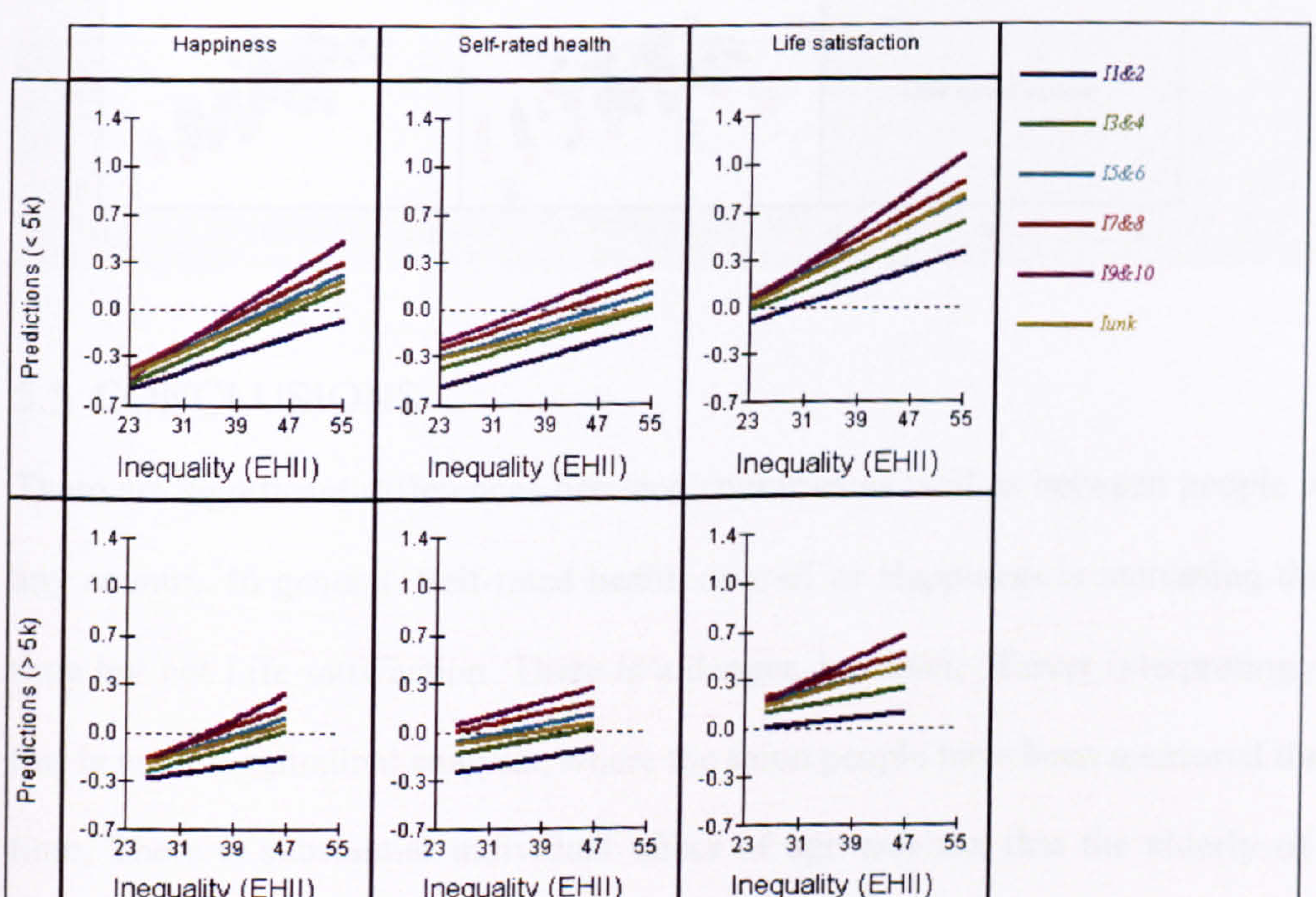
Figure 5.7 Separated outcome predictions associated with country income inequality



Income inequality of a country is associated generally with improved Happiness, Self-rated health and Life satisfaction but the effects are relatively shallow and have a wide confidence interval. It is maybe, however, that the posited relative inequality hypothesis works differentially for people of different individual income.

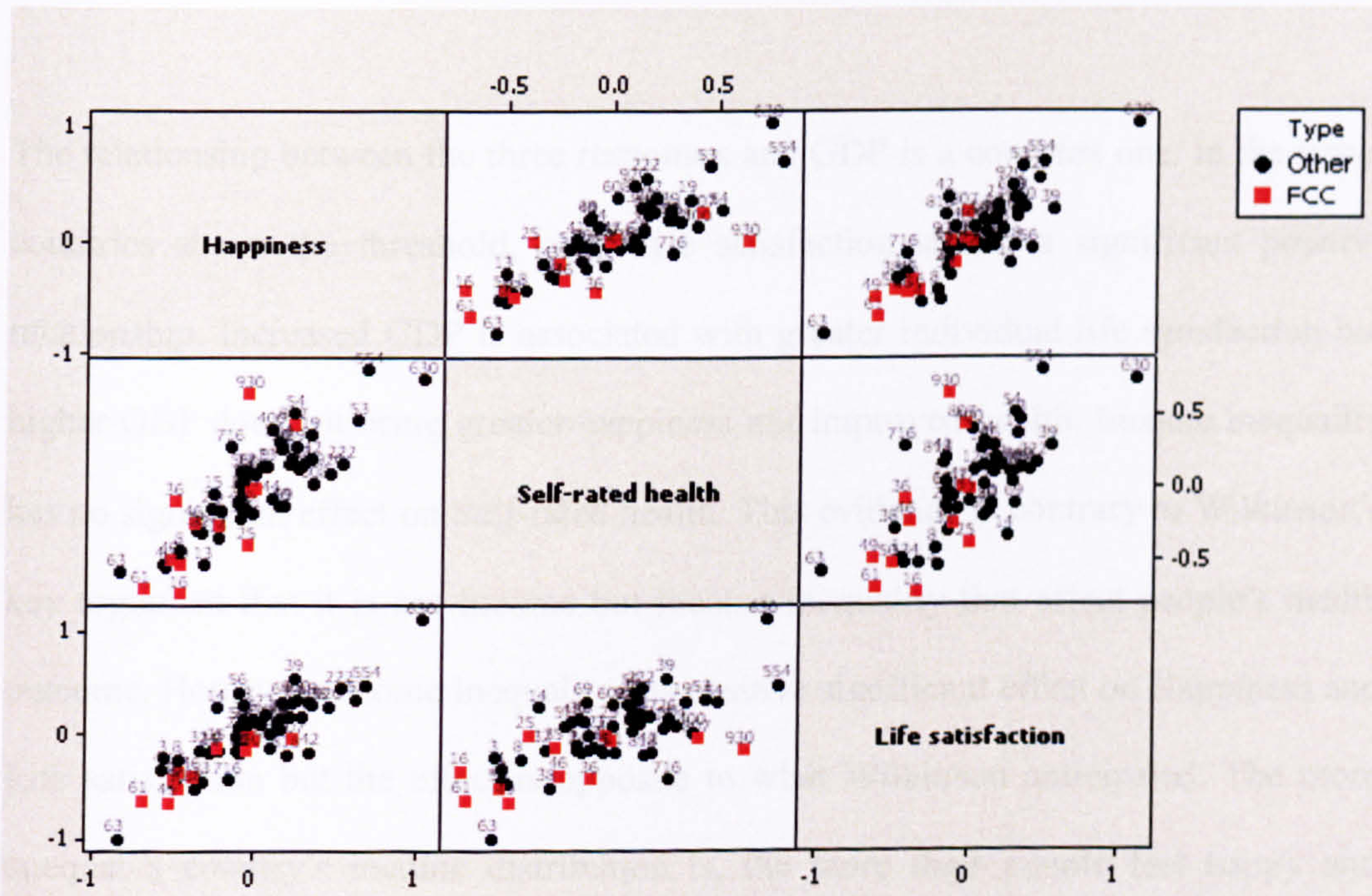
Consequently, the final Model 5 was fitted with interaction terms between individual income and inequality to allow testing this argument. Again the results are consistent, but do not support the Wilkinson hypothesis. All five groups of individual income show greater happiness, healthiness and life satisfaction as inequality increases, and this is the case both for countries above and below the threshold. The only new finding is there is a tendency, for Happiness and Life satisfaction to show an increasing gap between incomes as inequality increases. Rich people are even happier and more satisfied when there is high inequality; but, and against the Wilkinson hypothesis, there is still no sign of a negative effect of inequality.

Figure 5.8 Separated outcome predictions associated with country inequality for different individual income group of countries which are above (top row) / below (bottom row) five thousand US dollars in 1990



Turning now to the random between-country part of Model 5, the differences remain substantial and have not attenuated greatly as individual and country by wave variables have been included. Figure 5.9 shows the country estimated residuals against each other in a matrix plot. The novel and dramatic feature of this plot is the 'clumping' of the Former Communist Countries which all share the lower happiness, poor of health and lower life satisfaction.

Figure 5.9 Country differentials (based on Model 5) for all three outcomes



5.5 CONCLUSIONS

There are significant differences between countries as well as between people within any country. In general, Self-rated health as well as Happiness is increasing through time but not Life satisfaction. There is a danger, however, of over interpreting, since this is not a longitudinal analysis, where the same people have been measured through time. There is substantial individual effect of age and sex that the elderly of both genders tend to report poorer health and less happiness than young people, and live somewhat more satisfied lives. The separated, widowed and divorced (SWD), are more likely to report a worst situation than couples and singles; single people are less happy, have poorer health and less life satisfaction than couples. There is a very consistent 'dose-response' relationship between each of the three outcomes and individual income. Individual income affects the others to a similar degree and to a similar pattern (as income goes from the wealthiest quintile (9&10) to the lowest quintile (1&2), the less he/she tends to be happy, healthy and satisfied with himself/herself in life), with the largest effects being found for Life satisfaction.

The relationship between the three responses and GDP is a complex one. In the richer countries above the threshold, only Life satisfaction shows a significant positive relationship. Increased GDP is associated with greater individual life satisfaction but higher GDP does not bring greater happiness nor improved health. Income inequality has no significant effect on Self-rated health. This evidence is contrary to Wilkinson's key argument that it is not income but income inequality that affect people's health outcome. However, income inequality has positive significant effect on Happiness and Life satisfaction but the effect is opposite to what Wilkinson anticipated. The more unequal a country's income distribution is, the more their people feel happy and satisfied with their life.

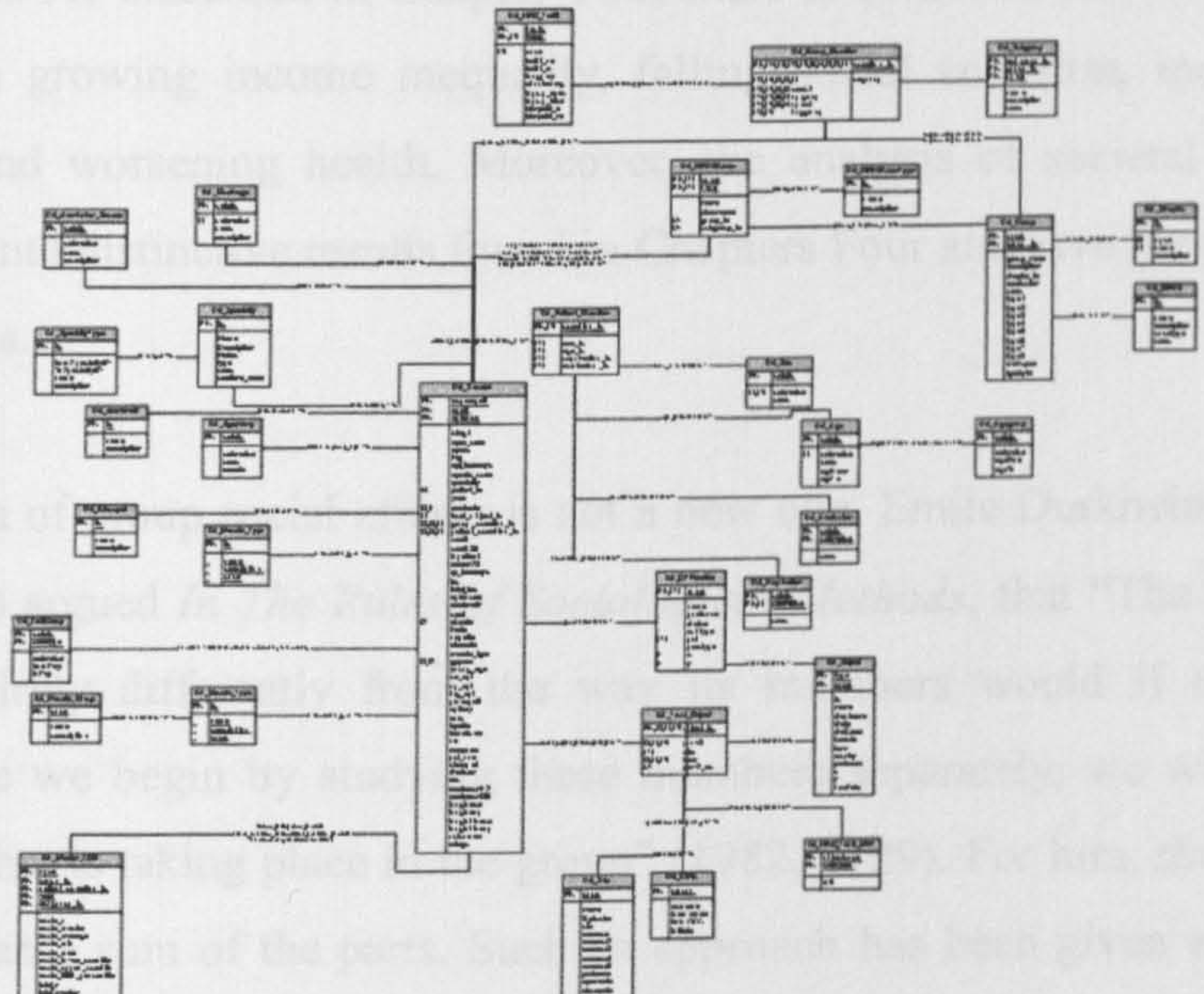
Wilkinson also stressed that people who have suffered the most from income inequality are poor people. I found that that while the income gap increased with inequality for Happiness and Life satisfaction, the improved response for the rich was not being achieved at the expense of a worse response for those with poorer individual income. There remains substantial between-country variation that is not accounted for by either individual characteristics nor GDP and inequality. It is a feature of this unexplained variation that the former Communist Countries all report lower happiness, poorer health and lower life satisfaction.

Chapter 6

An analysis of self-rated health and social trust using the World Values Survey

6.1 INTRODUCTION

This final analytical chapter analyses the relationships between self-rated health and individual and societal trust. As such, it extends the consideration of the Wilkinson relative income hypothesis into the realms of the effects of social capital and social cohesion. As discussed in Chapter Two, there is considerable recent interest in the links between growing income inequality, falling social cohesion, increasing psycho-social stress and worsening health. Moreover, an analysis of societal trust may give some insight into findings presented in Chapters Four and Five for the former Soviet bloc countries.



The idea of group social effects is not a new one. Emile Durkheim (originally published in 1895) argued in *The Rules of Sociological Method*, that "The group thinks, feels and acts entirely differently from the way its members would if they were isolated. If therefore we begin by studying these members separately, we will understand nothing about what is taking place in the group" (1982, p. 129). For him, the effect of the group is more than the sum of the parts. Such an approach has been given a major boost with the work of Robert Putnam through his development of the concept of social capital. Social capital is defined as the extent of connectedness and solidarity among groups in society which can be facilitated by social structure (such as trust, norms and networks). In *Bowling Alone* (1995), he argues that the USA has suffered a major collapse in civic, social, associational, and political life (social capital) since the 1960s, with serious negative consequences. His main finding is that virtually every traditional civic, social, and fraternal organization – typified by bowling leagues – had undergone a major drop in membership. Putnam makes a distinction between two kinds of social capital. *Bonding* capital occurs when you are socializing with people who are like you: same age, same race, same religion, and so on. But in order to create peaceful societies in heterogeneous multi-ethnic countries, one needs *bridging* capital. Bridging is what you do when you

Chapter 6

An analysis of self-rated health and social trust using the World Values Survey

6.1 INTRODUCTION

This final analytical chapter analyses the relationships between self-rated health and individual and societal trust. As such, it extends the consideration of the Wilkinson relative income hypothesis into the realms of the effects of social capital and social cohesion. As discussed in Chapter Two, there is considerable recent interest in the links between growing income inequality, falling social cohesion, increasing psycho-social stress and worsening health. Moreover, the analysis of societal trust may give some insight into distinctive results found in Chapters Four and Five for the former Soviet Bloc countries.

The idea of group social effects is not a new one. Emile Durkheim (originally published in 1895) argued *In The Rules of Sociological Methods*, that "The group think, feels and acts entirely differently from the way its members would if they were isolated. If therefore we begin by studying these members separately, we will understand nothing about what is taking place in the group" (1982, p129). For him, the effect of the group is more than a sum of the parts. Such an approach has been given a major boost with the work of Robert Putnam through his development of the concept of social capital. Social capital is defined as the extent of connectedness and solidarity among groups in society which can be facilitated by social structure (such as trust, norms, and sanctions). In *Bowling Alone* (1995), he argues that the USA has suffered a major collapse in civic, social, associational, and political life (social capital) since the 1960s, with serious negative consequences. His main finding is that virtually every traditional civic, social, and fraternal organization -- typified by bowling leagues -- had undergone a major drop in membership. Putnam makes a destinction between two kinds of social capital. *Bonding* capital occurs when you are socializing with people who are like you: same age, same race, same religion, and so on. But in order to create peaceful societies in heterogenous multi-ethnic countries, one needs *bridging* capital. Bridging is what you do when you make friends with people who are not like you, and you have trust in others.

There has been a growing interest in the effects of various forms of social capital on health and recent work has been summarised by Kawachi and Berkman (2000). An early example is Kawachi et al (1997) who demonstrated the ecological association at 39 states level in the USA between three different measures of social capital (social trust[1], other social capital measures[2] and group membership[3]) and income inequality and mortality in 1990. They found that a strong correlation of r = 0.77 between States with high levels of mistrust and high levels of all-cause mortality and 58% of the variance in total mortality is explained by the social trust. Other social capital measures shows nearly identical results with r = 0.79 from social trust, however, a weaker but still remain statistical significant correlation of r = -0.46 for the measure of group membership. They contend that income inequality has induced rising mortality rates via disinvestments in social capital. In another example, Lochner *et al.,* (2002) found a significant ecological association between low mortality rates and aggregated community level variations in interpersonal trust [4], reciprocity [5], and group membership [6] for 342 Chicago neighbourhoods in the USA.

Social capital in the form of trust has also been implicated in the mortality crisis being experienced by the former Soviet Bloc countries. For example, Kennedy et al (1998) argue that "Citizens living in societies with a high degree of social cohesion -- characterized by strong social networks and high levels of interpersonal trust -- seem to be healthier than those living in socially disorganized societies". Using household survey data they carried out a cross-sectional, aggregate analysis of the association between indicators of social capital and mortality rates across 40 regions of Russia. They found

---

[1] Social trust measured by responses to "Do you think most people would try to take advantage of you if they got the chance, or would they try to be fair?"
[2] Other social capital measured by agreement to "You can't be too careful in dealing with people" and "People mostly look out for themselves".
[3] Group membership measured by the per capita number of groups and associations to which residents in each state belonged.
[4] Perceptions of trust were measured by the proportion of residents in each neighborhood cluster answering strongly agree/agree to the question that "people in this neighborhood can be trusted"
[5] Reciprocity were assessed as the proportion of residents in each neighbourhood answering strongly agree/agree to the question that "people around here are willing to help their neighbours"
[6] Survey respondents were asked about membership in a variety of voluntary associations, including religious organizations, neighborhood associations, business or civic groups, neighborhood ethnic or nationality clubs, as well as neighborhood/local political organizations. From these responses, we constructed a measure of the average per capita associational membership in each neighborhood.

associations between indicators of social capital (mistrust in government, civic engagement in politics) and life expectancy, as well as mortality rates. They suggest that in the absence of civil society, more people in post-Soviet Russia rely on informal sources of support (friends, family) to deal with their day-to-day problems. Those lacking such sources of support may have been especially vulnerable to the economic hardships following the transformation to a market economy.

More recent studies have correctly moved away from aggregate analyses to multilevel approaches which allow the assessment of group or contextual effects in addition to, or interaction with, individual characteristics. A pertinent example for the present analysis is the study by Subramanian et al (2002) which investigated the effect of individual and group-level social trust on self-rated health for 40 US communities. They found that when community trust is low, there is little differential effect for individuals with a high and low trust. But at high levels of community trust, those expressing low trust experience much worse health. They describe (p531) their results as 'preliminary', suggesting that such complex interactions would benefit from more in-depth investigations. This is precisely what this chapter aims to do by replicating their analysis, using multilevel models and the WVS data to examine the relationship between individual self-rated health and social trust both at individual and country level, after taking account of individual demographic and income variables.

The aims of the present chapter are to examine and demonstrate:

- the overall relationship between individual self-rated health and individual demographic and economic factors

- the extent to which individual demographic and economic factors account for between-country variations in self-rated health;

- the overall relationship between individual social trust and self-rated health after taking account of individual demographic and economic factors;

- the extent to which individual social trust accounts for between-country variations in self-rated health;

- the overall relationship between country social trust (measured by aggregating individual responses to a question about interpersonal trust) and self-rated health;

- the extent to which this cross-level relationship between country trust and individual health is an artefact of individual perceptions social trust; Is there a contextual effect of country trust after taking account of individual differences in trust?

- Does between individual trust and contextual (country) trust interact in affecting health status; e.g. do people with high social trust living in a 'trusting' country feel a particular health advantage?

This list of questions represents a range of model fitted in exactly the same order as the original investigators, consequently the chapter will conclude with a comparison of the WVS results with that of Subramanian et al's 2002 paper.

The remainder of the chapter is in three parts; data, methodology and results with discussion. Firstly, in the data section, I outline the questions employed to operationalize ill-health and trust and consider how the WVS dataset needs to be manipulated to achieve the objectives outlined above briefly. Second, I will apply the technique of Binomial multilevel modelling. This is used here because the response variable to match the Subramanian et al paper is binary categories (good and poor health) thereby requiring binomial modelling; and the WVS data has a three-level structure individuals who are nested within different waves, with countries at the top of the hierarchical structure, thereby requiring multilevel modelling. In Chapter 4 we have introduced multilevel models for response variables with more than two categories; multinomial multilevel modelling, at length. The logistic models are simplifications that do not require detailed development. Consequently I will concentrate on the sequence of models fitted. This methodology allows us to model the micro level (age, sex and individual income), the macro level (mean income and income inequality) and the cross-level interactions between individual and country trust. The final part of the chapter reports a series of models of increasing complexity to evaluate the social trust variant of the Wilkinson hypothesis (1996).

6.2 DATA

The requirement was to replicate the Subramanian et al (2002) study and hence I need to use, as far as possible, the same concepts based on the same questions for both health and trust; the difference being that my places are countries as befits the original Wilkinson hypothesis rather than intra-national communities. This was indeed achieved with single exception of race.

6.2.1 The World Value Survey data

The analysis was based on scientific sampling of the World Values Surveys and European Values Surveys, which amalgamates over 60 surveys through four waves (1981, 1991, 1995 to 1997 and 1999-2001). This survey was compiled from a worldwide investigation of attitudes and norm on socio-cultural and political change by interviewing under the direction of Ronald Inglehart (1997, 2004) who is responsible for the cooperating, assembling and documenting the WVS with a group of other researchers.

6.2.1.1 Outcome measure

Self-reported overall health status of individuals has been commonly used as a valid indicator of health. As discussed in Chapter Four, there is an extensive amount of literature that shows this indicator is capable of assessing mortality risk independent of other medical, behavioural and psychosocial risk factors. It was determined from people's response to the following question: "How would you describe your state of health these days? Would you say it is excellent, very good, good, fair, or poor?". To facilitate comparable results with Subramanian et al (2002), I have reclassified the fivefold category to form a dichotomous outcome of self-rated health which 0 is for excellent, very good and good and 1 for fair and poor. In other words, we analyzed the underlying probability of reporting fair/poor health.

6.2.1.2 Independent variables

Self-rated health will be related to predictors at both the individual-level and country-level. At the individual level, we consider key demographic variables (age, gender, marital status) and income class characteristics. Perceptions of individual trust were determined by individual responses to a general question on interpersonal trust ("Generally speaking, would you say that most people can be trusted or that you can't be too careful in dealing with people?") with the potential responses being "people can be trusted," and "you can't be too careful,". At the country level, a contextual social trust variable, measured on a continuous scale, was derived by aggregating the individual responses to questions on interpersonal trust. Values were calculated by taking the arithmetic average of the weighted individual-level measures for each country at each wave to approximate national population values. In each country, the investigators were asked to provide a 4-digit weight variable to correct their sample to reflect national distributions of key variables. If no weighting was necessary, each case was simply weighted as 1. Given the relatively large size of the sample at each country for each wave (the minimum is 1000), there should not be a substantial problem of sampling reliability.

Figure 6.1 gives a general picture showing the changing average trust by all the country through waves. In general, average trust is more variable between countries from wave 2 to 3. Most west European countries are increasing from wave 1 to 2 (dark blue lines). The Nordic countries, shown by black lines, have the highest average level of trust. North and South America and Canada are decreasing in their average trust (Green lines). As expected, the former Soviet Bloc and its satellite states have the lowest average trust and getting wore (red lines). However, the worst trust of all is for Brazil.

Figure 6.1 Changing average trust: country & wave

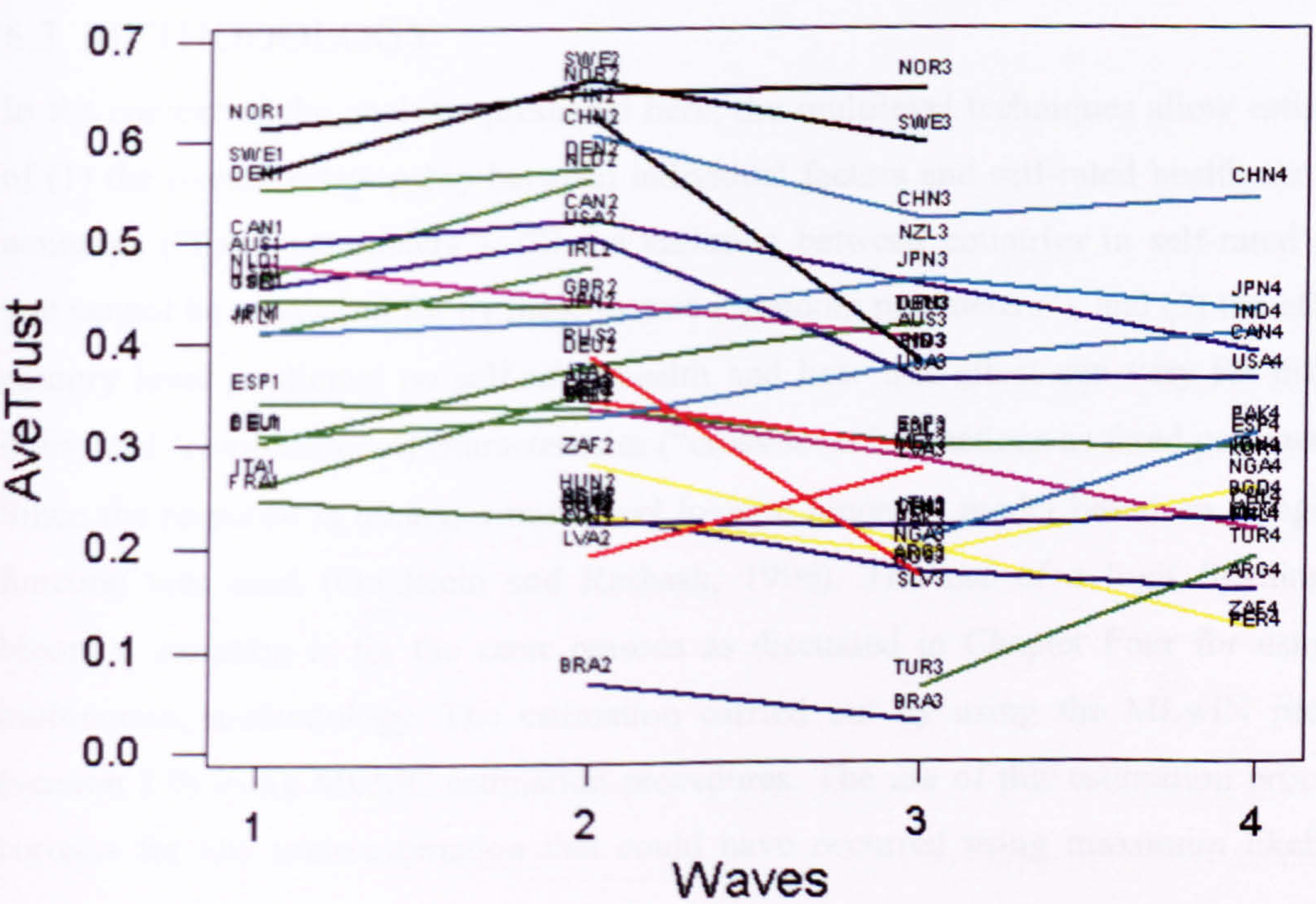


Table 6.1 shows the detailed mean social trust of former Soviet Bloc Countries and global average by wave. In general, global mean social trust decreases through waves as people trust each other less on average over time. There are no data for any country in wave one which used to be under communist government. All of the countries in each wave have lower social trust on average compared to global. The data show considerable variation, and this should allow a good test of the Trust hypothesis.

Table 6.1 The mean social trust for former Communist Countries compare to global by each wave

Wave	Country	Mean	Wave	Country	Mean	Wave	Country	Mean	Wave	Country	Mean
W1-W4	Global	0.3172									
W1	Global	0.4253	W2	Global	0.3546	W3	Global	0.2870	W4	Global	0.2563
			W2	Latvia	0.1905	W3	Russia	0.1696	W4	Moldova	0.1469
			W2	Slovenia	0.2110	W3	Moldova	0.1914	W4	Macedonia	0.2259
			W2	Bulgaria	0.2110	W3	Belarus	0.2051	W4	Albania	0.2259
			W2	Romania	0.2110	W3	Lithuania	0.2222			
			W2	Slovakia	0.2306	W3	Georgia	0.2255			
			W2	Hungary	0.2459	W3	Armenia	0.2255			
			W2	Poland	0.3448	W3	Azerbaijan	0.2255			
			W2	Russia	0.3841	W3	Latvia	0.2748			
						W3	Estonia	0.2748			
						W3	Ukraine	0.2748			
						W3	Poland	0.2940			

6.3 METHODOLOGY

In the context of the analysis presented here, the multilevel techniques allow estimation of (1) the overall relationship between individual factors and self-rated health across all countries ("fixed parameters"), (2) the variation between countries in self-rated health that cannot be accounted for by these factors ("random parameters"), and (3) the effect of country level predictors on self-rated health and how this effect can vary for different individual (compositional) characteristics ("cross-level interactions as fixed parameters"). Since the response is binary, a multilevel logistic binomial model based on a logit-link function was used (Goldstein and Rasbash, 1996). The use of a logit function and binomial variation is for the same reasons as discussed in Chapter Four for using the multinomial methodology. The estimation carried out by using the MLwiN program (version 2.0) using MCMC estimation procedures. The use of this estimation procedure corrects for any underestimation that could have occurred using maximum likelihood estimations (such as predictive/penalized quasi likelihood approximation) and allows the calculation of the Deviance Information Criterion (DIC) which provides a comparative measure of how well the model has been fitted. The smaller the value is, the better the model (Spiegelhalter et al, 2002). All models were estimated using logit (logarithm of the odds) function but for ease of interpretation, and as appropriate, I use logits transformed to proportions, odds ratios (OR) or both. All models are fitted with accompanying sample weights at level 1 to adjust for potential sampling bias in the survey.

The following first six models (Model 1 to Model 4B) were sequentially developed in the same manner as Subramanian et al (2002). The models are:

Model 1: a three-level null (empty) model of individuals (level 1) nested within 4 waves (level 2) nested within 69 countries (level 3) with no predictor variables in the fixed and random parts except as set of dummies representing waves (with wave 1 as reference) in the fixed of the model with sample weights on. This model provided a baseline for the comparison of the degree that the compositional and contextual variables account for the variation in self-rated health between countries or within country between waves in subsequent models.

Model 2: this is the same as Model 1, but includes all the individual predictors (except individual trust) in the fixed part of the model. The model assessed the effect of individual predictors on self-rated poor health. Individual predictors were entered in the model in two sequential steps: first, the demographic variables age, sex, age*sex interaction and marital status was included (Model 2A) and then socioeconomic status variables (individual income groups) were added (Model 2B). The contextual variation in self-rated poor health between countries was estimated before and after taking into account the compositional effect of individual demographic and economic variables.

Model 3: this the same as Model 2, but considers the fixed effect of country- aggregated social trust on individual self-rated poor health and the extent to which it explains the country-level differences.

Model 4 A&B: this is the same as Model 3, but considers the effect of interpersonal trust at the individual level to evaluate the relative importance of individual-level versus country-level social trust (Model 4A). In addition, we also considered how the effect of country social trust on self-rated poor health differed for low- and high-trust individuals (Model 4B).

The final most complex model 4B is specified as follows as below:

$$\text{Log}\ (\pi_{ijk}) = \beta_{0jk}X_{0ijk} + \beta_1 W2_{jk} + \beta_2 W3_{jk} + \beta_3 W4_{jk} + \beta_4 F_{ijk} + \beta_5 Age_{ijk} + \beta_6 F{*}Age_{ijk} + \beta_7 Mswd_{ijk} + \beta_8 Msig_{ijk} + \beta_9 Munk_{ijk} + \beta_{10} I1\&2_{ijk} + \beta_{11} I3\&4_{ijk} + \beta_{12} I7\&8_{ijk} + \beta_{13} I9\&10_{ijk} + \beta_{14} Iunk_{ijk} + \beta_{15} HTrustI_{ijk} + \beta_{16} TrustC_{jk} + \beta_{17} HTrust{*}TrustC_{ijk} + (\mu_{0jk}X_0 + v_{0k}X_0)$$

Where the variables are defined as follows:

- $\text{Log}\ (\pi_{ijk})$ is the log of the odds of reporting poor health for individual i in wave j in country k at time point j for individual i;
- X_{0ijk} is the constant;
- $W2_{jk}$, $W3_{jk}$ and $W4_{jk}$ are dummy variables (with W1 as the base) for each survey wave;
- F_{ijk} is a indicator variable identifying females with a 1, males with a 0;
- Age_{ijk} is the age for each individual centred around age 40;
- $F{*}Age_{ijk}$ is interaction between the female dummy and continuous age;
- $Mswd_{ijk}$, $Msig_{ijk}$ and $Munk_{ijk}$ (with Couple as the base) represent dummy variable of each marital status category; widowed/separate/divorced, single and unknown respond, respectively;
- $I1\&2_{ijk}$, $I3\&4_{ijk}$, $I7\&8_{ijk}$, $I9\&10_{ijk}$ and $Iunk_{ijk}$ (with I5&6 as the base) are dummies for income quintiles;

- $HTrustI_{ijk}$ is the dummy variable for individual level interpersonal trust with low trust as the base;

- $TrustC_{jk}$ is a continuous scale of national aggregated trust centred around 0.3; and finally

- $HTrust*TrustC_{ijk}$ is the interaction between individual and country social trust.

The fixed parameters which show the overall relationship between self-rated health and both individual and country predictors consequently have the following meaning:

β_0	is the global mean log-odds of poor health of the reference group (a middle income 40 years old married male with low interpersonal trust who live in a country with the average trust around 0.3) in 1981;
$\beta_1, \beta_2, \beta_3$	define the increasing or decreasing differential log-odds of reporting poor health in each survey time point in comparison to 1981;
β_4	is the gender gap for self-rated poor health; the difference between male and females in the log-odds;
β_5	estimates the change in the log-odds of poor self-reported health for a change in one year of age for female;
β_6	the interaction level between age and gender to see whether the differences between gender change with age;
$\beta_7, \beta_8, \beta_9$	define the difference for marital status in contrast to the base of 'couple';

$\beta_{10}, \beta_{11}, \beta_{12}, \beta_{13}, \beta_{14}$ is the differentials for each income group in contrast to middle income 5 and 6 on the quintile scale;

β_{15} defines the differential effect for high individual social trust compare to low social trust;

β_{16} is the estimate of the effect of country trust on log-odds of being in poor health;

β_{17} estimates the effect of cross-level interaction between country and individual social trust;

The random part summarises the variations between countries and waves in self-rated health that cannot be accounted for by the included factors. There is a complex expression for the variability around the global trends that is differentiated for low and high individual social trust ($v_{0k} X_0 + \mu_{0jk} X_0$) where

v_{0k} is the country differential in 1981 for low trust

μ_{0jk} is the differential between waves within each country

6.4 RESULTS

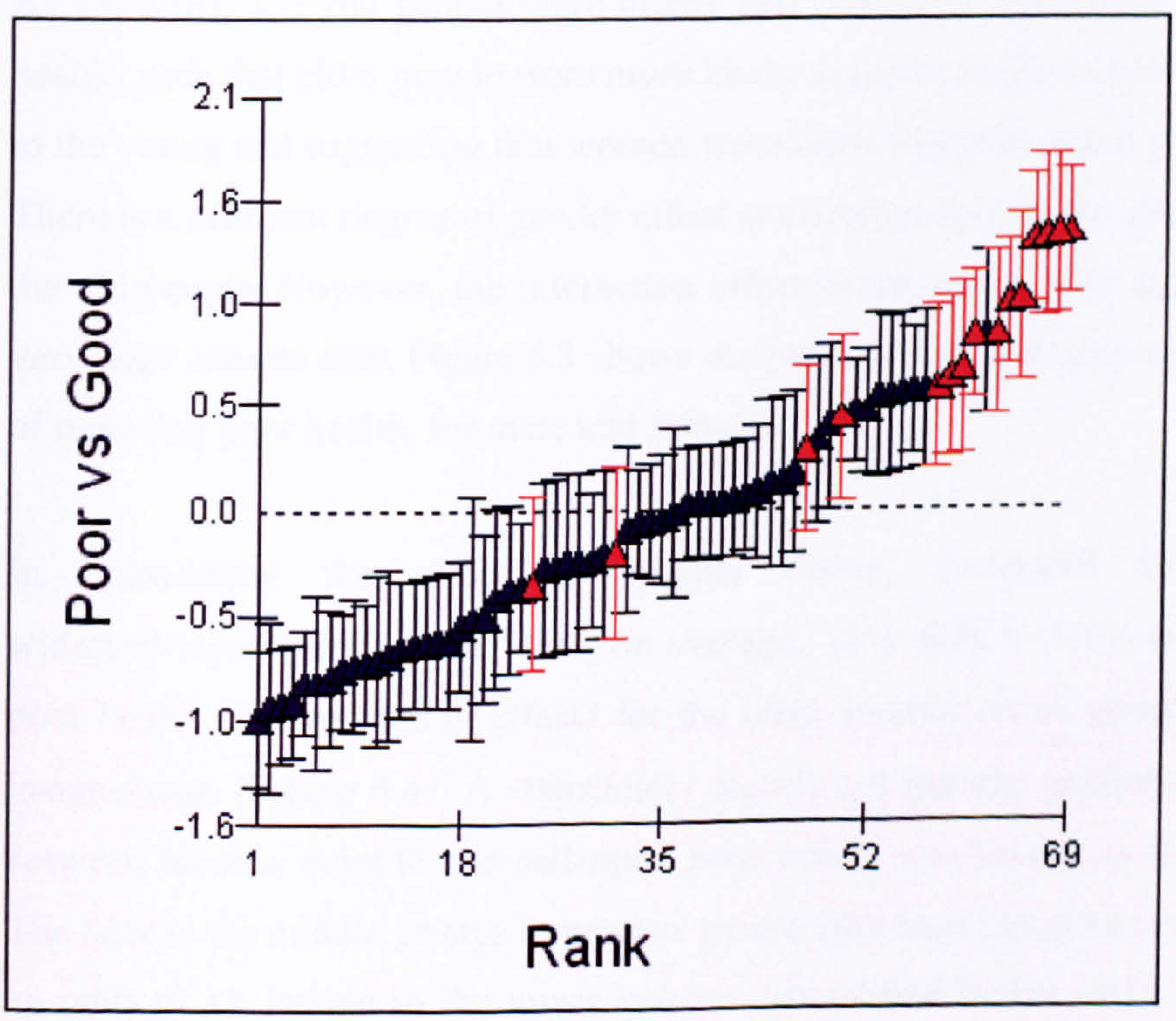
Table 6.1 provides a summary of the final data considered for the analysis. Except for age, individual characteristics were specified as categorical variables, with a base and a set of contrast indicator dummies. The total number of individual observations from 69 countries was 171,264. After excluding the missing data on the predictor variables, I conducted a multilevel regression analysis on 163,328 individuals nested within 4 waves (1981, 1991, 1995-97, 1999-2000) at level 2 for 69 countries at level 3 (Table 6.2).

The results of the multilevel models are presented in Tables 6.3 with their estimate of standard errors and the DIC statistics in the order discussed above. We found that approximately 46.7% of the respondents in 1981 reported to be in poor/fair health by converting the logit estimates from model 1 (the null model). However, through time the probability reduced to 19.1% in 2000. Gradually, less people said they are in poor health, but we have to be careful for this interpretation again, due to different countries participating in different waves.

Figure 6.2 shows the between-country variation as point estimates on a logit scale, for poor health in comparison to good. These are values derived across all the waves of the survey. In this figure, a 'caterpillar plot' shows the logit estimates and their 95% confidence intervals plotted against rank with former Soviet States highlighted in red. The horizontal dashed line indicates the average of all countries set to zero. Quite a few countries are significantly above and below the average at both ends statistically. Countries which reported low poor health in logit are mainly west wealthy economies (Ireland, Switzerland and Canada etc.) and Nordic countries (such as Denmark, Norway, Finland and Sweden). Nine out of ten countries which reported high poor health are former Soviet States. Consequently, the average logit proportion reporting poor/fair health in a country is very different from each other.

Figure 6.2 Between-country differences (based on Model 1)

Logit and 95% confidence intervals against rank of country



Turning now to the level 2 variances of within countries between waves, it can be seen that these are much smaller ($\sigma^2_{\mu 0}$ = 0.038) than between countries (σ^2_{v0} = 0.473). Thus, there was some variation within countries across waves, but the effects are not as substantial as between-country differences. Those countries with poor health tended to remain so across time.

Model 2 additionally includes a range of individual variables for respondents, representing demographic characteristics (Model 2A) and income status (Model 2B). The reference group in model 2B – the model with all the individual predictors – is a 40-year-old, married male who belongs to the middle income category (I5&6). The independent differential effects (compared to the above reference group) of each individual variable are now presented. The main effect for each of the covariates was estimated after controlling for the remaining ones (see Table 6.3, model 2B). There is a substantial improvement as witnessed by the substantial lowering of the DIC (13426.99,) It is the sum of 11584.36 in the model 2A plus a further reduction 1842.63 in the Model 2B).

As expected, age and gender were highly and positively associated with poor self-rated health, such that older people were more likely to report being in poor health as compared to the young and suggesting that women were more likely to report poor health than men. There is a different degree of gender effect at different ages as the gender gap is larger for the old people. However, the interaction effect is not statistically significant, despite the very large sample size. Figure 6.3 shows the predicted relationship between age and odds of reporting poor health, for male and females.

In considering the effect of marital status, compared to married people, widowed/separated/divorced were, on average, 19% (OR = 1.19) more likely to report poor health. The remaining effects for the other marital status groups were smaller and insignificant (Figure 6.4). A statistically significant inverse gradient in the relationship between income quintile and self-rated poor health was found, as shown in Figure 6.5. The base is the middle groups in income groups (the base category consequently received an odds of 1). People in the lower income groups had higher odds of poor health with group 1&2, the most disadvantaged group, having an OR of 1.71 and that for group 3&4 being 1.26. In contrast, the higher household income group, 7&8 had an OR of 0.82 while group 9&10, the most advantage of all, had the lowest OR at 0.68. There is therefore a clear and non-linear dose response relationship between individual income and self-rated health.

Figure 6.3 The estimated relationship between age and relative odds of reporting poor health, for male and female at different ages (based on Model 2B)

Figure 6.4 The estimated relationship between marital status and relative odds of reporting poor health based on Model 2B

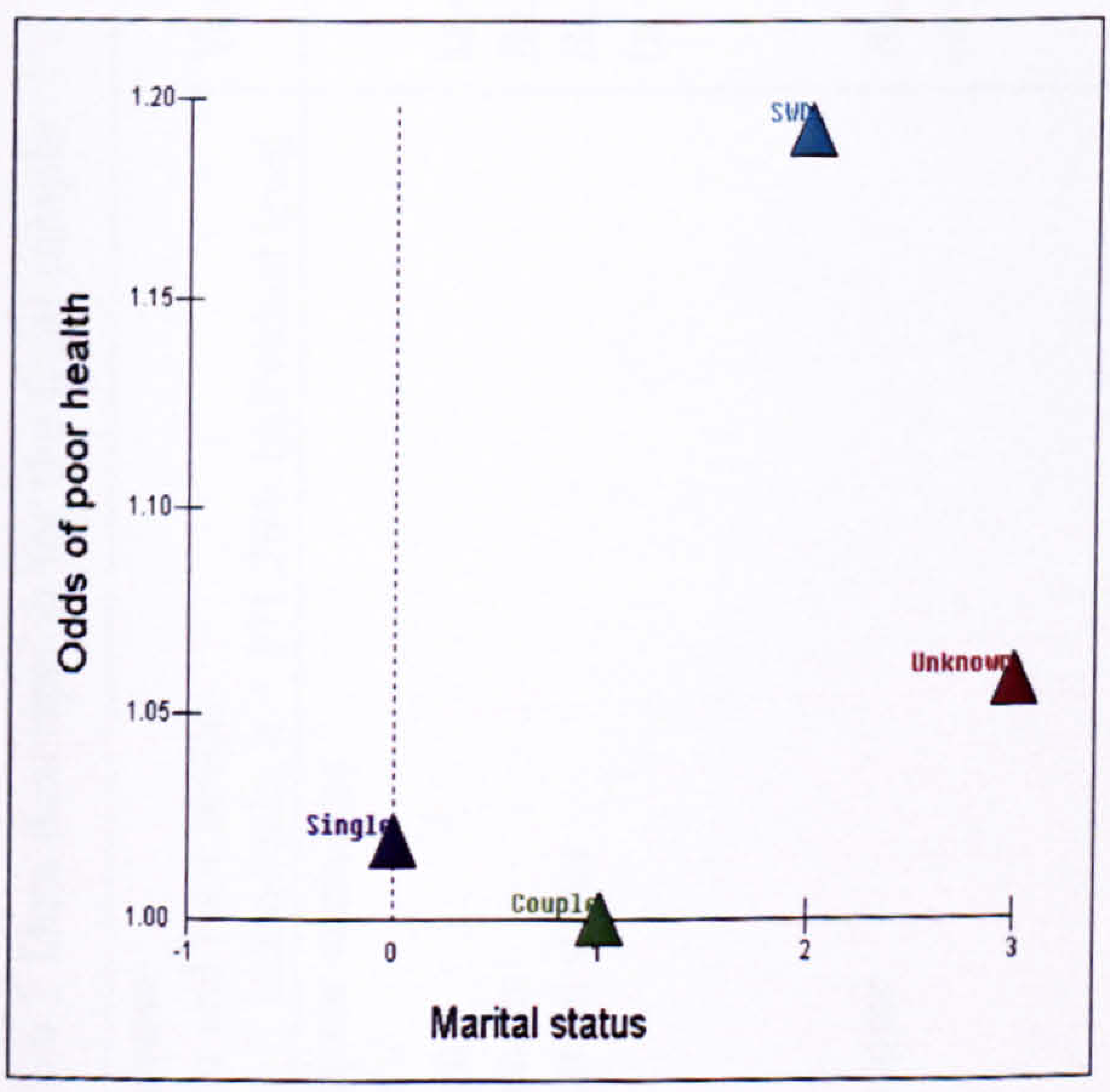


Table 6.2 Data description for the final sample

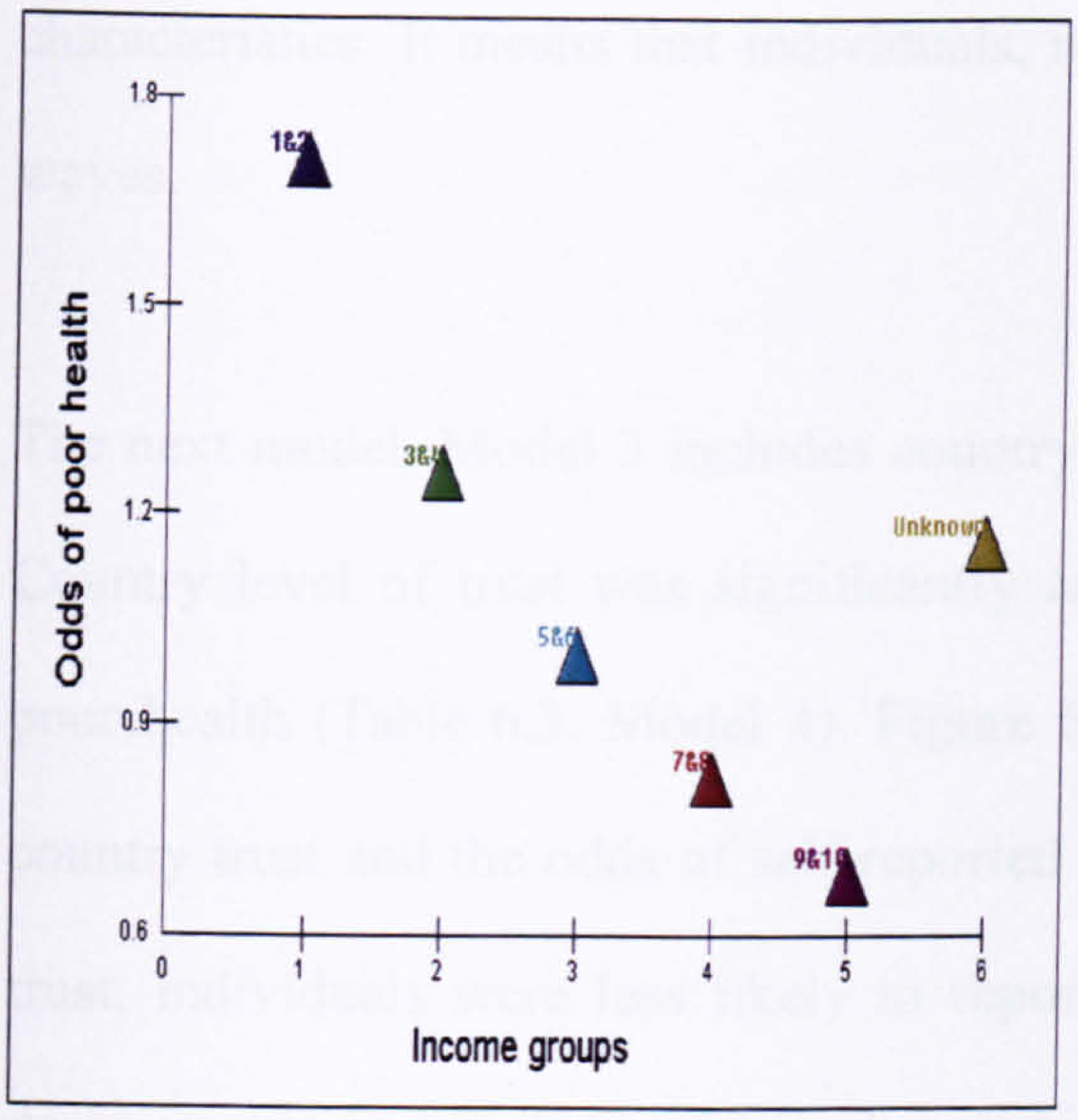
Response		
Fair and Poor health	Yes (n = 59,483, 36.4%)	No (n =100,953, 61.8%)
Level 1, individuals, n = 171,264: individual level		
Predictor variables		
Level 1		
Age	Mean = 40 years	Range = 15-97 years
Gender	*Base: Male* (n = 79,382, 48.6%)	Contrast: Female (n = 83,946, 51.4%)
Marital status	*Base: Married* (n = 106,710, 65.3%)	Contrast: Single (n = 38,739, 23.7%) Separated/Widowed/Divorced (n = 17,589, 10.8%) UNK (n = 290, 0.2%)
Income	*Base: 5&6* (n = 34,534, 21.1%)	Contrast: 1&2 (n = 31,962, 19.6%) 3&4 (n = 40,377, 24.7%) 7&8 (n = 23,554, 14.4%) 9&10 (n = 13,433, 8.3%) UNK (n = 19,468, 11.9%)
Individual social trust	*Base: Low trust* (n = 113,101, 69.2%)	Contrast: High trust (n = 50,227, 30.8%)
Level 2		
Wave, n = 116 country*wave	*Base: Wave 1* (n = 19,543, 12.3%)	Contrast: Wave 2 (n = 46,171, 28.3%) Wave 3 (n = 51,423, 31.5%) Wave 4 (n = 46,191, 28.3%)
Level 3		
Countries, n = 69: country-level		
Aggregated social trust	Mean = 0.308	Range = 0.03 to 0.66

Table 6.3 Fixed and random part results for the Binomial analytical models (in logits)

Fixed Parameters	Model 1 (S.E.)	Model 2a (S.E.)	Model 2b (S.E.)	Model 3 (S.E.)	Model 4a (S.E.)	Model 4b (S.E.)
Constant	-0.331 (0.106)	-0.474 (0.110)	-0.530 (0.112)	-0.536 (0.106)	-0.432 (0.106)	-0.432 (0.106)
Wave 1 (base)						
Wave 2	-0.045 (0.071)*	-0.138 (0.068)	-0.193 (0.073)	-0.133 (0.075)	-0.131 (0.075)	-0.129 (0.075)
Wave 3	-0.186 (0.080)	-0.300 (0.076)	-0.342 (0.081)	-0.373 (0.082)	-0.375 (0.083)	-0.374 (0.083)
Wave 4	-0.387 (0.087)	-0.522 (0.083)	-0.543 (0.089)	-0.586 (0.090)	-0.587 (0.090)	-0.586 (0.090)
Individual predictors						
Female		0.287 (0.012)	0.268 (0.012)	0.268 (0.012)	0.269 (0.012)	0.269 (0.012)
Age (centered around 40)		0.037 (0.001)	0.035 (0.001)	0.035 (0.001)	0.035 (0.001)	0.035 (0.001)
Female*Age		0.001 (0.001)*	0.001 (0.001)*	0.001 (0.001)*	0.001 (0.001)*	0.001 (0.001)*
Marital status						
Couple (base)						
Widowed/ Separate/Divorced		0.279 (0.019)	0.176 (0.020)	0.176 (0.020)	0.171 (0.020)	0.171 (0.020)
Single		0.065 (0.016)	0.019 (0.016)*	0.019 (0.016)*	0.022 (0.016)*	0.022 (0.016)*
Unknown		0.083 (0.140)*	0.056 (0.140)*	0.057 (0.140)*	0.052 (0.141)*	0.050 (0.141)*
Income groups						
5&6 (base)						
1&2			0.536 (0.019)	0.535 (0.019)	0.526 (0.019)	0.526 (0.019)
3&4			0.230 (0.017)	0.230 (0.017)	0.225 (0.032)	0.225 (0.017)
7&8			-0.193 (0.021)	-0.192 (0.021)	-0.187 (0.021)	-0.186 (0.021)
9&10			-0.380 (0.027)	-0.379 (0.027)	-0.360 (0.027)	-0.359 (0.027)
Unknown			0.148 (0.022)	0.149 (0.022)	0.140 (0.022)	0.140 (0.022)
Social capital						
High trust					-0.357 (0.013)	-0.350 (0.014)
Country predictor						
Trust				-1.624 (0.446)	-1.338 (0.446)	-1.260 (0.448)
Individual/Country interaction						
High trust*Trust						-0.234 (0.104)
Random parameters						
Level 3	0.473 (0.086)	0.559 (0.100)	0.538 (0.097)	0.442 (0.081)	0.441 (0.081)	0.440 (0.081)
Between countries						
Covariance						
Individual low trust						
Level 2	0.038 (0.008)	0.034 (0.008)	0.039 (0.009)	0.040 (0.009)	0.040 (0.009)	0.040 (0.009)
Within country between waves						
Covariance						
Individual high trust						
Deviance information criterion	195294.98	183710.62	181867.99	181861.19	181104.22	181098.57
Change in DIC from previous model	--	11584.36	1842 63	6.80	763.04	5.65

All estimates are significant at 0.05 probability level, except those marked by *, which have a probability greater than 0.05.

Figure 6.5 The estimated relationship between income groups and relative odds of reporting poor health based on model 2B

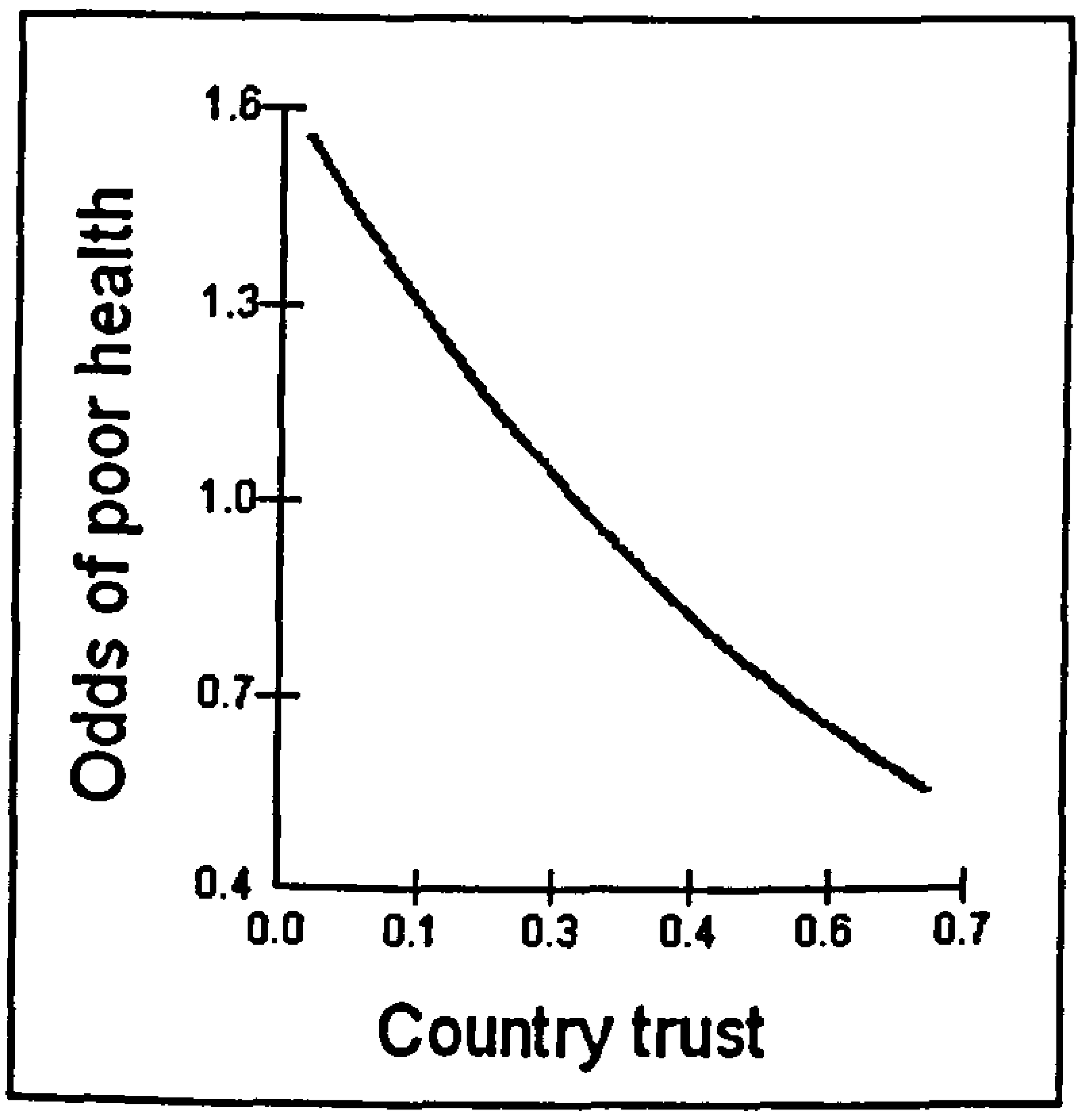


The estimated effect of these individual variables mentioned above reflects the same result from Chapter Four that young comparing to old people, women to men, and widowed/separated/divorced to married people are more likely to report poor or fair health. The degree of individual income effect to their health is nonlinear. The inclusion of the set of income variables in model 2B as compared to model 2A produces a very substantial reduction in the DIC showing the importance of individual incomes association with self-reported health, but we cannot, with repeated cross sectional data, rule out the reciprocal camal path that poor health leads to poor income.

It is noticeable that the inclusion of demographic variables in model 2A increases the estimate of between-country differences and that they remain substantial and significant when individual income variables are included. The substantial between-country variations strongly suggest that individual differences cannot explain away the country differences. The much smaller within-country between-wave variations barely changed (0.038 to 0.039) after taking accounts all these individual characteristics. It means that individuals, not countries, have great similarity across waves.

The next model, Model 3 includes country levels of trust as a contextual predictor. Country level of trust was significantly and negatively associated with self-rated poor health (Table 6.3, Model 3). Figure 6.6 shows a plot of the relation between country trust and the odds of self-reported health. In countries with higher level of trust, individuals were less likely to report poor health after controlling for their demographic and socioeconomic characteristics. The inclusion of this variable also accounts for some of the between-country variance, which falls from 0.538 to 0.442 but this remains statistically significant. And again a smaller DIC resulted then in the last model (by 6.80) showing that including country-level trust on a wave basis leads to an improved model with better fit

Figure 6.6 The estimated relationship between countries aggregated social trust and relative odds of reporting poor health based on Model 3



In Model 4 the main contextual effect of country trust remains largely the same after taking account of individual differences in trust perception, suggesting that the aggregate community trust effect is not an artefact of individual perceptions of social trust. However, when explored further, a significant cross-level interaction effect between country trust and individual social trust was observed in model 4B.

Compare to Model 3, Models 4A & 4B are much better fits according to the reduction in the value of DIC of 763.04 and 5.65, respectively. Figure 6.7 a) plots the predicted relationship between country trust and the odds of reporting poor health, for low and high-trust individuals based on results from model 4B. As can be seen, there is what is known as a 'consensual effect', in that individual and contextual trust are both important and the sign goes in the same direction: if you have high individual trust and live in a country with a high trust you tend to be healthier.

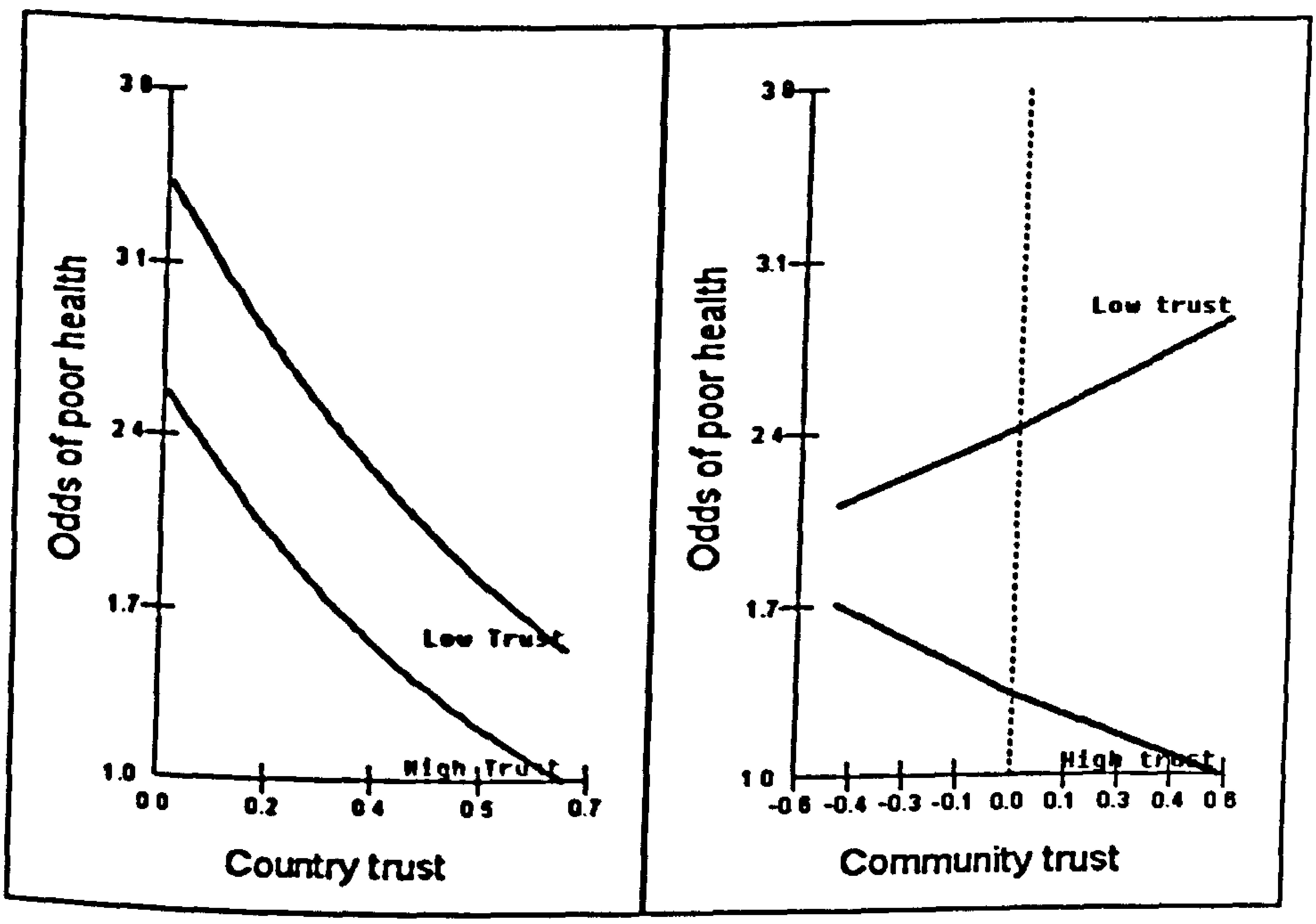
The major findings are compared with Subramanian et al's in Figure 6.7. The left-hand graph shows the relationship of the WVS, given in terms of odds, between poor self-rated health and society social trust (measured as the proportion of the country at a particular wave who answer 'yes' to the question concerning whether you can trust people who live in the same place). Cross-level interactions are fitted, but are not found to be substantial. There is not the same complex relationship of the right-hand side of the diagram (which is drawn using the same scale on the vertical axis) as was found by Subramanian and his colleagues.

At the world scale low trust at both the individual and country level are mutually reinforcing in being related to poor heath. The effects of societal trust are moderate ones with a twofold difference in poor health between the extremes of trust. But they do give some support to the 'social capital hypothesis' that individual and societal cohesion are related to individual health.

Figure 6.7 The comparison result for WVS analysis to those of Subramanian et al: cross-level interactions between individual trust and place trust

a) WVS

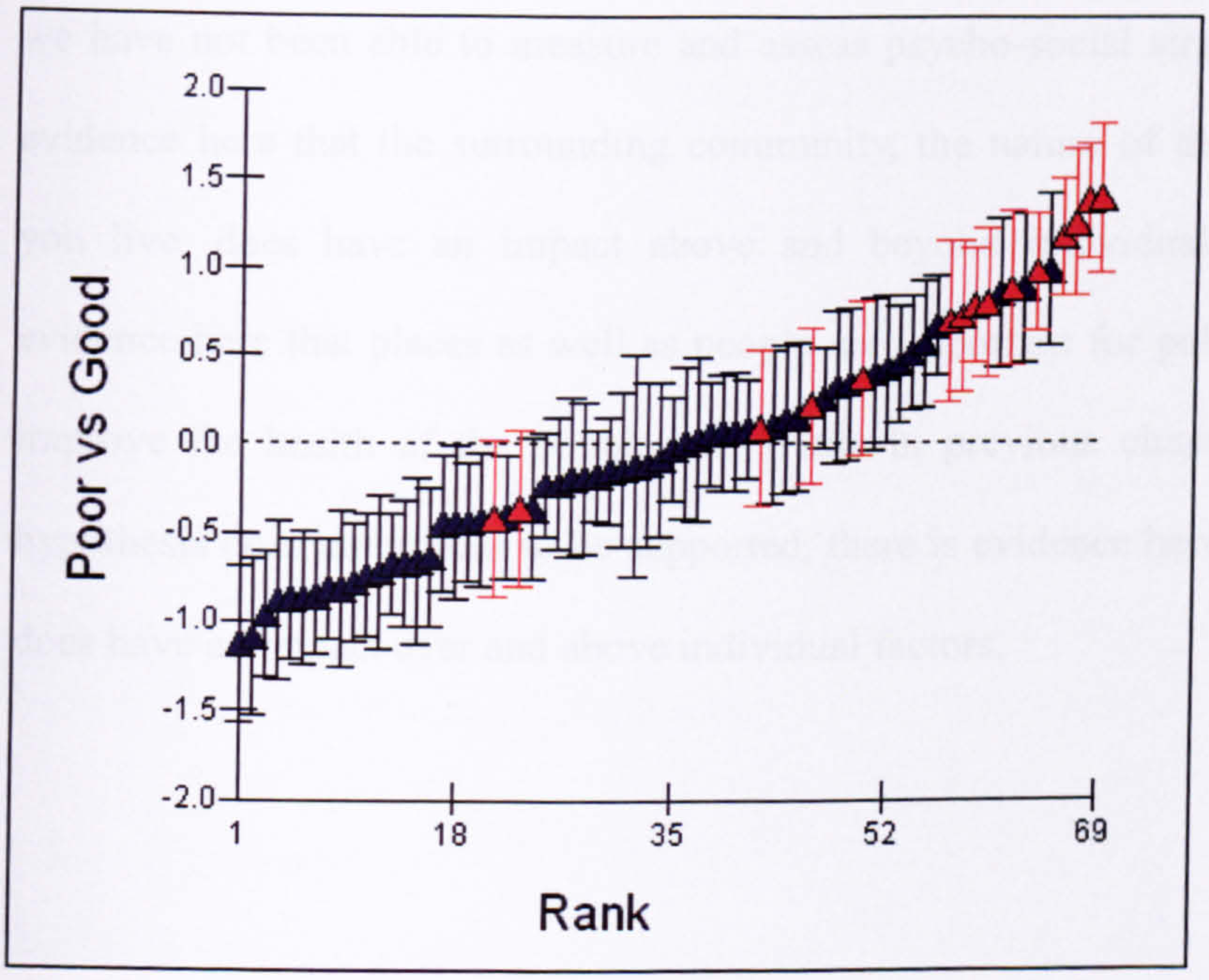
b) Subramanian et al



Moreover, Figure 6.8 shows the between-country variation of Model 4. Here seven out of top ten reporting poor health countries are the former Soviet States. The Soviet effect in terms of a revised Figure 6.2 has diminished. In other words, individual and mean social trust matter for people from Soviet Union countries to report themselves in poor health, after taking account of demographical variables and individual income.

Figure 6.8 Between country differences (base on Model 4B)

Logit and 95% confidence intervals against rank of country



6.5 CONCLUSIONS

In this study, I have found a contextual effect of country trust on health after taking account of individual socio-democratic, income variables and individual social trust. Social trust both at the country level and at the individual level have a significant, if modest, impact on people's perception of their health.

People tend to have poorer health when they have low levels of trust and when the nation they live in has a lower overall trust; a consensual effect. This contrasts markedly with the strong negative cross-level interaction between individual and community level in the study of Subramanian et al, which those authors find difficulty in explaining. Indeed, there is a major problem of the design of that study in that 'community' is very ill defined and varies from a single municipalities to entire States in the USA. Here, with the WVS data, the results are more plausible. If you do not trust others and the country you live in has a low trust of others, a lack of bridging capital in Putnam's terms, you are more likely to report poor health. While we have not been able to measure and assess psycho-social stress directly there is evidence here that the surrounding community, the nature of the society in which you live, does have an impact above and beyond individual factors. There is evidence here that places as well as people are important for policy intervention to improve the health of the population. While in previous chapters the inequality hypothesis does not appear to be supported, there is evidence here that societal trust does have an impact over and above individual factors.

Chapter 7
Conclusions

7.1 Introduction

This final chapter is in three parts. The first section summarises what we have learnt from the empirical modelling in relation to the Wilkinson's hypothesis. The second part addresses the difficulties of using survey data collected internationally on health, income and subjective well being. The final part considers some further limitations of the research and how the work may be developed in the future to tackle these limitations.

7.2 Summary

In this thesis, health outcomes at the aggregated level (life expectancy at birth derived from World Bank database) as well as at the individual level (self-rated health from the World Values Survey) datasets have been investigated to evaluate Richard Wilkinson's relative income hypothesis. This states that it is the degree of inequality of income within a developed society rather than the absolute income that is the main determining factor for health. For him, the key mechanism causing health inequalities is not material disadvantage but psycho-social stress arising from perceived relative disadvantage. The distinctive feature of this thesis is that recent methodological developments in random coefficient modelling have been applied to empirically evaluate the relative income hypothesis at the scale it was originally proposed, that is countries.

7.2.1 Aggregate level

Chapter Three applied the methodology of group trajectory modelling using Nagin and Jones's PROC TRAJ software. The models fitted to country-level aggregated data for the last thirty years show that there are distinctive trajectories of life expectancy at birth, with the identified groups of countries have a clear and marked geographical distribution. According to the Bayesian information criteria, the World can be classified into ten groups as the best solution with each of the groups having different life expectancy patterns through the study period. However, the four-group solution accounts for most of the complexity of the data structure (see Table 3.12).

In brief, Model 4 (Figure 3.19), shows that most Sub-Sahara Africa has a pattern of 'worst life expectancy (LE)' from 50 years in 1970 which increases to 55 in the late 1980s but then declines substantially to 40 in 2002. This is mainly because of the spread of HIV/AIDS that has shortened people's lives. With slightly better life expectancy at 51 years in 1970 are the countries of North Africa, they have a dramatically improved pattern to 62 years by 2002. Another group of countries includes all of the former Soviet States and some countries in the central Africa. It has a 'steady LE' trajectory which has remained at about 60 years throughout the period. These countries have not enjoyed the improvements that have been seen elsewhere. This reflects the unstable situation politically and economically after the mid 1980s in these places. Most western economies as well as Australia, New Zealand and China have been grouped together with the 'best LE' during the study period that increases from 63 years in 1970 to 69 by 2002. This latter group of countries are the ones that Wilkinson suggests conform to his theory that while GDP is not important, income inequality is a key determinant of how long people live.

A key finding is that the effect of income and income inequality are very different from group to group of countries as shown in Table 3.14. For worst LE countries, neither GDP nor inequality affects life expectancy significantly although they have relatively larger positive and negative impact respectively on this group than other groups. Both income and inequality positively affect life expectancy in 'improving LE' countries but inequality is not statistically significant. For 'steady LE' countries, GDP is more important in predicting group membership than inequality, and it has a positive and significant impact on life expectancy. In contrast, inequality is more important than GDP in countries having the 'best LE'. In summary, it is necessary to study countries separately in sub-groups for analysing the mechanisms of how income and income inequality influence people's health. Using aggregate data at the country level and examining changing life expectancy over a thirty year period, there is indeed support for the Wilkinson hypothesis in the countries with the best life expectancy.

7.2.2 Individual level

The original Wilkinson work was based on data aggregated to the country level. Such an approach is always open to the charge of the ecological fallacy in that the observed relationships may have been induced by data aggregation. Gravelle's (1998) paper gave the critique a new twist by mathematically showing that a non-linear relationship between a response and a predictor at the individual level induces a relationship between the mean of the response and the *variance* of the predictor when the data are aggregated. That is Gravelle suggested that a non-linear relationship between individual health outcomes and individual income, will produce an artefactual relationship between self-rated health and income inequality when the unit of analysis is a country.

Chapter Two of this thesis clearly demonstrated that it was possible to simulate data that produced the results predicted by Gravelle. Moreover, it was shown in that chapter that multilevel modelling could be used to identify substantive ecological effects and that this procedure could model simultaneously at the micro level of individuals and at the macro level of countries. Consequently it was vital to test the relative income hypothesis using individual level data on health outcomes and country income inequality having already included individual income (allowed to take a non-linear form) in the model. Chapters Four, Five and Six analyse the World Values Survey as the key test of the relative income hypothesis. Wilkinson also notes that if a country has been through the economic transition (above five thousand US dollars in 1990), GDP is no longer the factor affecting people' health. The World Values Survey which includes a wide range of countries above and below the 'threshold' provides an excellent test bed for this part of the hypothesis.

7.2.2.1 Self-rated health and income inequality: a multilevel analysis

In Chapter Four, a marked non-linear relationship has been found between self-rated health and individual income after taking account of individual demographic variables. As postulated by Gravelle, this relationship is indeed a non-linear one. The effect of GDP on health is found to be negative and insignificant in both countries above and below the threshold after taking account of individual demographic and income variables. In other words, the greater a country's income, the more people in the country report less poor/fair health, but these effects are statistically insignificant.

Moreover, and in marked contrast to Wilkinson's work, income inequality has a negative effect for countries above and below the five thousand US dollars threshold, for reporting poor/fair health. The more unequal the income distribution a country has, the less people report that they are in poor/fair health. It is the opposite effect anticipated by Wilkinson. Moreover, it is statistically significant for countries above five thousand dollars while being insignificant for those below.

In this global study, there are still substantial between-countries differences which have not been explained away by both compositional (income, age, marital status and gender) and contextual variables (GDP, income inequality) in the models. Moreover, there is a tendency of reporting poor health for the former Soviet States, and these are distinct from other countries. It can be concluded that there is not enough evidence supporting the Wilkinson hypothesis once individual's income and its differential impact are taken into account on self-rated health. GDP has little effect in both rich and poor countries and income inequality does not have the postulated effect on morbidity. Even when interactions are allowed between individual income and inequality, poor people in the most unequal counties do not appear to experience the worsened poor health predicted by Wilkinson.

7.2.2.2 Multivariate modelling subjective well being

In Chapter Five, attention focuses on the Wilkinson hypothesis and subjective well-being. He has argued consistently that effects of relative income operate not through material pathways, but through psycho-social mechanisms. People in unequal countries are unhappy, dis-satisfied with their life and are consequently ill. This element of the hypothesis is analysed using a multivariate multilevel model and the World Values Survey.

I find a substantial individual income effect consistent with a 'dose-response' relationship between each of the three outcomes and individual income categories that goes from the wealthiest quintile (9&10) to the lowest quintile (1&2), such that the lower the income, the less he/she tends to be happy, healthy and satisfied with himself/herself in life, with the largest effects being found for Life satisfaction. When the time-varying country level variable for GDP is included in the model, the results are against Wilkinson's argument that further increases in country's income for the countries above threshold is no longer associated with increases in Self-rated health. This also applies to Happiness. However, there is significant positive relationship between GDP and Life satisfaction. For the countries below the threshold, the more a country's income the less they report good health and happiness but people are more satisfied with life.

In terms of country-level income inequality, there is no significant effect on self-rated health (thereby and unsurprisingly duplicating the findings in Chapter Four), but importantly more unequal distributions of income affect Happiness and Life satisfaction in the opposite direction to that proposed by Wilkinson. Wilkinson also mentioned that people who have suffered the most from income inequality are poor people. In the study, poor people suffered lower life satisfaction and were less happy and there is a large difference of happiness and life satisfaction in the most unequal income distribution country. The former Communist Countries all share lower happiness, poorer health and lower life satisfaction. Again when individual income is included in the model, the postulated relationship between inequality individual health, happiness and satisfaction is not found and this is the case for poor and rich individuals.

7.2.2.3 Social trust: a comparison study with the USA

Chapter Six focuses on a variant of the Wilkinson's hypothesis that suggests people will be unhealthy when there is a lack of social cohesion in a country. This approach may explain the poor self-rated health of the former Soviet satellites, but it is crucial that individual measures of social cohesion are included in the model as well a measure for the country as a whole. This chapter replicates the study of Subramanian *et al.*, (2002) but unlike that study which used US communities, countries are used as the meaningful higher-level unit.

The result shows that health has been affected by country trust (a contextual effect) after controlling for individual socio-democratic, income variables and individual social trust.

There is a significant and positive effect on health by social trust at the country level as well as at the individual level with the same direction for both variables: increased trust is associated with better health. In other words, people tend to report themselves to have better health if the level of trust within the country and between each other is high.

This finding conflicts with the USA study published by Subramanian *et al.*, (2002) that found a strong negative cross-level interaction between individual and community level, which they found difficult to explain. This may be because there is a major fault in their study that the definition of places varies from single municipalities to entire States. It is important for that policy intervention focuses on places as well as people to improve people's health. Although there is no support for the inequality hypothesis in Chapter Four and Five, social trust at the country has influenced the above individual factors. In the review Table 2.2, none of the papers investigates social trust in the country level and there is no consistent relationship between individual health and individual and community social trust. However, Subramanian et al., (2002) argues that while the income inequality hypothesis is not always replicated; social trust does get more sustained support. The present study adds further evidence to this argument for the importance of social trust as both an individual and contextual variable.

7.3 Limitations and Discussion

7.3.1 Validity and reliability

Chapters Four to Six are based on the World Value Survey. With such large scale cross-national studies, it is very easy to raise questions about validity and reliability. For instance, do the variables really represent true differences in health or are they merely cultural constructs. There are well-known difficulties and errors associated with cross-cultural surveys in many aspects of the design, such as the questions, the sampling, the translations and the interviewing techniques.

The validity and reliability limitation in cross-cultural comparability in survey research can be minimized with carefully designed questionnaires and carefully worded and constructed questions and Inglehart (2004) who is the project leader argues that this has been done for the WVS. Esmer (2004) has investigated and commented on the four basic aspects of sample survey in the WVS: questionnaire design, sampling, data collection and analysis. Firstly, the understanding for the same question or even a phrase is very different from nations to nations, family to family and person to person. For example, a nuclear family in Norway refers to a very small one while in Nigeria it may mean a whole village. However, no matter what the exact meaning of the word, the 'family' is very important for people universally. In the last wave of the World Value Survey, they found 89.4% said "very important". Only 1.1% and 0.3% answered that family was "not important" or "not important at all" for them, respectively. This gives very strong evidence that people have a common concept in all societies despite the different definitions. It is difficult to think of another plausible explanation for this extremely low variance in the responses from almost one hundred thousand respondents scattered all over the World. The same argument can be extended and applied to self-rated health as well as other aspects of well-being.

Secondly, if a construct is appropriately measured it should be relatively stable over time. The gap between each wave of the World Values Surveys is approximately five years. The basic values and demographic structure in a society should consequently remain fairly stable over such a short period of time, unless there is a traumatic event like a war or severe economic crisis. Esmer in 2004 showed an example using data of respondent's perspective to a question on their personal health from the third (1995-97) and fourth (1999-2001) waves of the WVS, to argue that the responses have a strong cultural component that does not change much over time. Suppose a Nigerian and Japanese have the same health status from a medical point of view, are unlikely to report the same health status subjectively. If this is true and subjective health is a reasonably enduring cultural trait, a comparison of the distributions should give some clues to the adequacy of given samples.

Table 7.1 compares the percentages of respondents who said the state of their health was good or very good in the two waves of surveys, for the 21 WVS countries which data are available. The results are promising in that in 16 out of 21 countries in this table, the discrepancy is five percentage points or less, that is within acceptable sampling error limits. Nigeria and South Africa show larger shifts from one wave to the next. There could be a substantial number of explanations for these significant differences. Nevertheless, 19 out of 21 countries show differences of eight points or less between the two waves and the mean shift from one wave to the next is only 3 percents point.

Thirdly, a sampling strategy called "probability sampling with quotas" by Sudman (1976) has been applied in the World Values Survey, which according to him, "these quotas are normally determined for the smallest geographic area for which information is available". In statistical terms, it is believed that the error sum of squares is only a small proportion of the total sum of squares. Overall, there are a large number of references discussing the problem of validity in cross-cultural research in the WVS. It has been concluded that despite the very real and very serious difficulties that critics of cross-cultural research make, the data shows a high degree of cross-cultural validity (Esmer, 2004).

Table 7.1 Percent of respondents describing their health as good or very good

	1999-01	1995-97	Difference
Albania	75	73	2
Argentina	65	62	3
Bosnia Herzegovina	64	62	3
Chile	69	61	8
China	61	68	-7
India	63	59	4
Japan	55	56	-1
S Korea	78	76	2
Mexico	61	56	5
Moldova	30	31	-1
Nigeria	90	75	15
Pakistan	66	64	2
Peru	49	50	-1
Philippines	57	52	5
Puerto Rico	73	68	5
S Africa	78	68	10
Spain	76	68	8
Turkey	64	61	3
Macedonia	71	73	-2
USA	84	79	5
Serbia-Montenegro	56	51	5
	Mean difference		3.4

Source: Esmer, (WVS, 2004, p397)

7.4 Future work

This final section briefly spells out what future work is required for further investigating and developing Wilkinson's hypothesis, and thereby attempts to address limitations of the current work. There are four broad areas for future work:

7.4.1 True panels

First, it is necessary to have a longitudinal study for causal work as with repeated cross-sectional studies such as the WVS, it is impossible to tell if low individual income is really a consequent of poor health or a cause of it. While in general the reverse causation argument on downward social mobility is not strongly supported (Espen, 1996), the case for a causal relationship is made most powerfully when a change in the cause, be it between income or inequality, precedes a change in the health outcomes. The possibilities of carrying out such work could be based on the European Community Household Panel Survey (ECHP, http://forum.europa.eu.int/irc/dsis/echpanel/info/data/information.html) in which the same individuals are followed over time but they live in different countries. Such a study could focus on changing individual income and changing subjective health. The data on individual income is also likely to be of a much higher quality than the WVS.

7.4.2 Epidemiological transitions

Second, Wilkinson (1996) claims that when countries pass through the income threshold to a developed economy, they also pass through the epidemiologic transition. That is the main causes of diseases changes from communicable to chronic disease in which income inequality is a key determinant. To evaluate this part of the hypothesis requires a study of changing of a wider set of health outcomes such as cause-specific mortality and an engagement with epidemiological theory (Omran, 1971). One possibility would be to look at a country such as Taiwan which (in addition to being the home country of the candidate) has been the subject of a Wilkinson-like aggregate analysis of mortality (Chiang, 1999) and it has a number of features that allow a quasi-experimental approach (Moon and Gould, 2000).

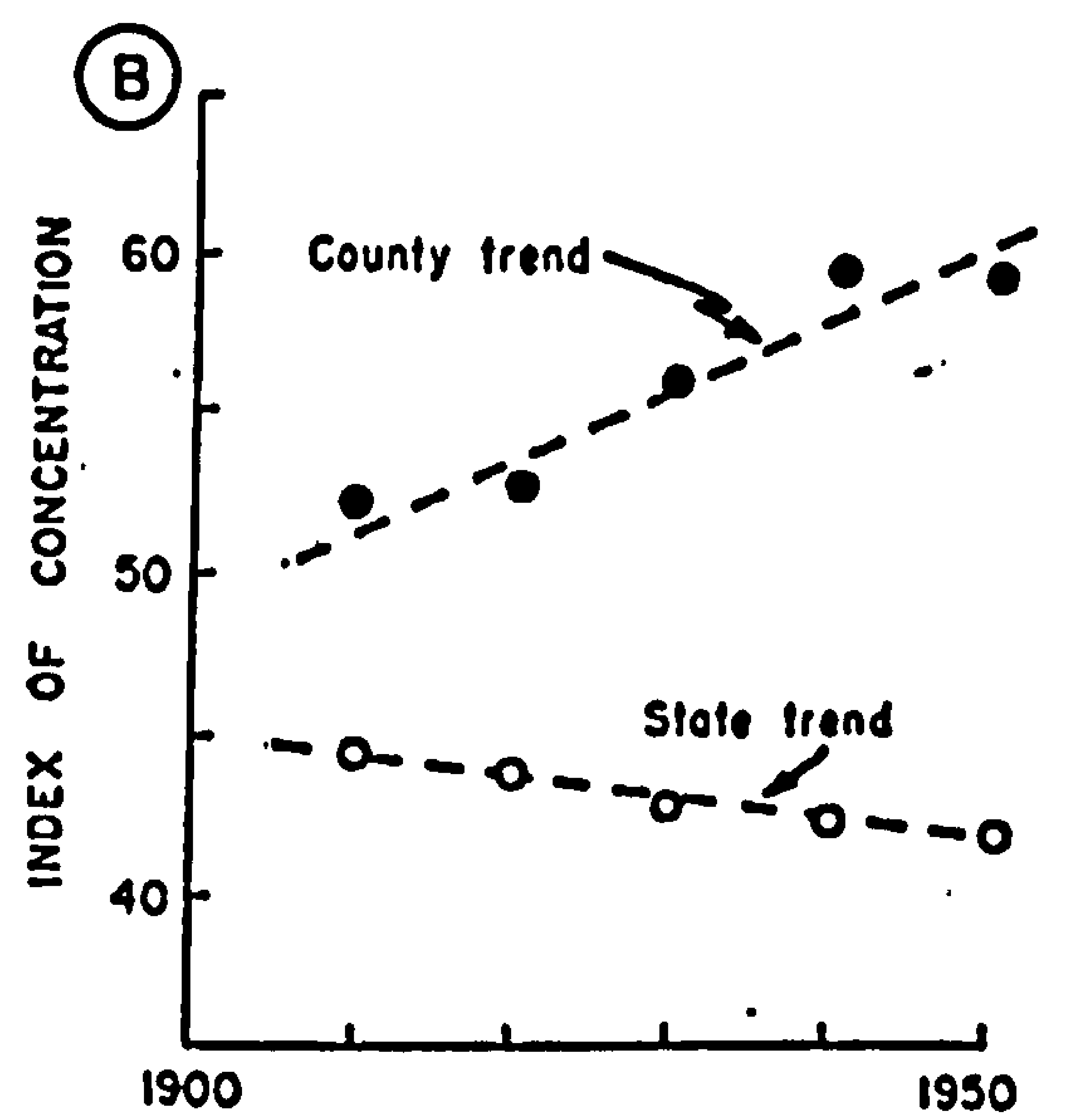
In the 1950's Taiwan was a poor agricultural country with a gross national product per capita of no more than US$200 (£72 valued at 1950s prices) but by 1995 the country an exportoriented economy with a figure of US$12,396 (£7853 valued at 1995 prices). Moreover, the economic development of Taiwan has been accompanied with improved income inequality. According to Chiang (1999), the Gini coefficient for Taiwan has decreased from 0.56 in 1953 to 0.28 in 1976 and gradually increased to 0.32 in 1995. Individual death records, some 3 million deaths with area codes are available for 1971-2001. The Global Burden of Disease (1990) classification of deaths, which divides causes of death at the broadest level into three groups (non-communicable diseases in Group 1, injuries in Group 2, and Communicable diseases in Group 3) could be used to look at the changing proportion of deaths in each category. Taiwan is therefore an ideal 'laboratory' to examine changing patterns of mortality as a country undergoes rapid socio-economic change. Unfortunately socio-economic data is only available for areas and not people, but these data would allow the assessment that of the epidemiological transition aspect of the Wilkinson hypothesis.

7.4.3 Scale

Third, scale is an important issue in the study of change as it is possible to have differential patterns at different scales occurring simultaneously (Haggett, 1965, see Figure 7.1 which shows the American population concentrating at the finer spatial scale and dispersing at the wider State scale over a 50 year period). Moreover, income inequality is clearly a scale-related concept. Thus in a study of a city, there could be gross inequalities between neighbourhoods, but equality within a neighbourhood. Consequently if a study is done at a neighbourhood scale it may give very different results than if undertaken at a city-region scale. In the main, the large literature on the Wilkinson hypothesis does not discuss this issue although he himself uses nation-states to stand for societies, but he also points approvingly to studies done at the regional level, such as Kawachi and Kennedy and colleague's analysis in 1997 at the State level in the USA.

One of the exceptions to this general failure to discuss scale is the study by Soobader and LeClere in 1999. They suggest that current approaches to the effect of income inequality on morbidity and mortality ignore the essentially geographic nature of that inequality. They also argue that both of unit of analysis (ecologic versus individual) and the level of aggregation may, in fact, be critical to the nature of the relationship between income inequality and mortality. Inconsistent findings across units of analysis and levels of aggregation may result not from a flaw in the basic premise that income inequality affects health outcomes but rather that all of the etiologic pathways have yet to be explored at a variety of scales. In their study of counties (coarser scale) versus tracts (finer scale), they found no effects of income inequality on morbidity at the latter scale, because these areas were essentially homogeneous. This is because residential segregation associated with income inequality is a process that generates areas of the city with internal similarity, but is different from other areas of cities.

Figure7.1 Scale-specific results: indices of population concentration for the United States by countries and states



Source: Haggett (1965, p201)

One way of approaching this is issue if scale is through using the more detailed geographies that are available in the WVS surveys, but some countries have detailed codes (Iceland) but for others (UK) the degree of detail is limited. A better possibility is to be able to vary the size of the neighbourhoods as Johnston and others (1997) have done for data derived from the British Household Panel Survey to form 'bespoke' neighbourhoods.

7.4.4 Social trust

Fourth, the key finding is that social trust is an important factor that affects health in this study. However, the cross-sectional nature of WVS makes it difficult to find the casual relationship between individual health and individual and aggregated level (country) social trust. Moreover, social trust is only one of the mechanisms (others such as networks of relationships, reciprocity and social norms etc.) that generated social capital/social cohesion. As have been reviewed in Table 2.2 of Chapter Two, there is no consistent finding to support the social capital hypothesis that affects health in different scale of data. This is because different indices have been applied for each study to measure social capital/social cohesion of that society. There is a need to study a wide range of these social capital aspects as possible.

Therefore, an alternative longitudinal data set: the European Community Household Panel Survey, which has mentioned above is ideal for us to test social capital hypothesis of a range of nations. This data set not only has variables of individual income, health, and their detail demographics but also a wide range covering of social capital aspects such as membership of associations, frequency of taking to neighbours, friendship networks, how people obtained employment (through various means including information from friends and relatives). Another important point is that weighting will always be open to dispute in forming the contextual variable. Thus, we need to ask what is the best way to sum indicator variables of social capital or how to weight at each scale into an overall index. Such work will need to be driven theoretically and empirically, but the findings of Chapter Six suggest that such work should be actively followed-up.

REFERENCES

Adler, N.E., Boyce, W.T., Chesney, M.A., Folkman, S. and Syme S.L. (1993) Socioeconomic inequalities in health: no easy solution. *Journal of the American Medical Association*, 269: 3140-3145.

Alderson, M.R. (1988) *Mortality, morbidity and health statistics*. Macmillan: Basingstoke.

Ben-Shlomo, Y., White, I.R., and Marmot, M. (1996) Does the variation in the socioeconomic characteristics of an area affect mortality? *British Medical Journal*, 312: 1013-1014.

Blakely, T.A., Kennedy, B.P., and Glass, R. et al. (2000) What is the lag time between income inequality and health status? *Journal of Epidemiology and Community Health*, 54: 318-319.

Blakely, T.A., Lochner, K., and Kawachi, I. (2002) Metropolitan area income inequality and self rated health-a multilevel study. *Social Science and Medicine*, 54: 65-77.

Blakely, T.A., O'Dea, D., and Atkinson, J. (2003) No association of income inequality with adult mortality within New Zealand. *Journal of Epidemiology and Community Health*, 57: 279-284.

Brame, R., Nagin, D.S. and Wasserman, L. (Forthcoming) Exploring some analytical characteristics of finite mixture models. *Journal of Quantitative Criminology*.

Brief, A., Butcher, A., George, J., and Link, K. (1993) Integrating bottom-up and top-down theories of subjective well-being: The case of health. *Journal of Personality and Social Psychology*, 64: 646-653.

Browne, W.J. (2004) *MCMC estimation in MLWiN.* Institute of Education, London.

Browning, C.R., and Cagney, K.A. (2002) Collective efficacy and health: neighbourhood social capital and self-rated physical functioning in an urban setting. *Journal of Health Social Behaviour*, 43: 383-399.

Bryk, A.S., and Raudenbush, S.W. (1987) Application of hierarchical linear models to assessing change. *Psychological Bulletin*, 101: 147-158.

Bryk, A.S., and Raudenbush, S.W. (1992) *Hierarchical linear models: applications and data analysis methods.* Newbury Park, CA: Sage

Catalan, R. (1998) Book review special of Richard Wilkinson's unhealthy societies: the afflictions of inequality: an epidemiological perspective. *Journal of Community Applied Social Psychology*, 8: 165-168.

Cattell, V. (2001) Poor people, poor places, and poor health: the mediating role of social networks and social capital. *Social Science and Medicine*, 52: 1501-1516.

Chiang, T.L. (1999) Economic transition and changing relation between income inequality and mortality in Taiwan: regression analysis. *British Medicinal Journal*, 319: 1162-1165.

Cleveland, W.S. (1979) Robust locally weighted regression and smoothing scatterplots. *Journal of Statistical Association*, 74: 829-836.

Coburn, D. (2000) Income inequality, social cohesion and the health status of populations: the role of neo-liberalism. *Social Science and Medicine*, 51: 135-146.

D'Unger, A.V., Land, K.C., McCall P.L., and Nagin, D.S. (1998) How many latent classes of delinquent/criminal careers? Result from mixed Poisson regression analyses. *The American Journal of Sociology*, 103: 1593-1630.

Deininger, K., and Squire, L. (1996) A new data set measuring income inequality. *World Bank Economics Review*, 10: 569-591.

Diez-Roux, A.V., Link, B.G., and Northridge, M.E. (2000) A multilevel analysis of income inequality and cardiovascular disease risk factors. *Social Science and Medicine*, 50: 673-687.

Doyal, L. (2001) Sex, gender, and health: the need for a new approach. *British Medical Journal*, 323: 1061-1063.

Drukker, M., Kaplan, C., Feron, F., and Os, J.V. (2003) Children's health-related quality of life, neighbourhood socioeconomic deprivation and social capital. A contextual analysis. *Social Science and Medicine*, 7: 825-841.

Drukker, M., SBukaa, S.L., Kaplan, C., McKenzie, K., and Os, J.V. (2005) Social capital and young adolescents' perceived health in different sociocultural settings. *Social Science and Medicine*, 61: 185-198.

Durkheim, E. (1895, 1982) *The rules of sociological method*, ed. Lukes, S. (tr. Halls, W.D., 1938). New York: Free Press.

Elgar, F.J., Roberts, C., Parry-Langdon, N., and Boyce, W. (2005) Income inequality and alcohol use: a multilevel analysis of drinking and drunkenness in adolescents in 34 countries. *European Journal of Public Health*, 15: 245-250.

Esmer, Y. (2004) Cross-cultural comparisons, survey methodology and the values surveys. In *Human beliefs and values: a cross-cultural sourcebook based on the 1999-2002 values surveys.* Inglehart, R. et al., Eds. Siglo XXI Editores. Maxico.

Espen, D. (1996) Social mobility and health: cause or effect? *British Medical Journal*, 313: 435-436.

Firebaugh, G. (2003) *The new geography of global income inequality.* Harvard University Press.

Fiscella, K., and Franks, P. (2000) Individual income inequality, health and mortality: what are the relationship? *Health Service Research*, 35: 307-318.

Fisher, K.J., and Li, F. (2004) A community-based walking trial to improve neighbourhood quality of life in older adults: a multilevel analysis. *Annals of Behavioural*

Medicine, 28: 186-194.

Franzini, L., Caughy, M., Spears, W., and Eugenia, M.E.F. (2005) Neighbourhood economic conditions, social processes, and self-rated health in low-income neighbourhoods in Texas: a multilevel latent variables model. *Social Science and Medicine*, 61: 1135-1150.

Galbraith J.K., and Kum H. (2002) Inequality and globalization: judging the data. (http://www.brookings.edu/gs/research/projects/glig/WorldBankJune2002Web.ppt)

Galbraith, J.K., and Kum H. (2004) Estimating the inequality of household incomes: a statistical approach to the creation of a dense and consistent global data set. UTIP working paper No. 22.

Galbraith, J.K., and Kum, H. (2003) Estimating the inequality of household incomes: filling gaps and fixing problems in Deininger and Squire. UTIP working paper No.22. (http://utip.gov.utexas.edu/web/workingpaper/UTIP22-REVISED.pdf)

Ghosh J.K., and Sen, P.K. (1985) On the asymptotic performance of the log likelihood ratio statistic for the mixture model and related results. *Proceedings of the Berkeley Conference*, Volume 11.

Gilks, W.R., Richardson, S., and Spiegelhalter, D.J. (1996) *Markov chain Monte Carlo in practice. Interdisciplinary statistics.* London, Chapman and Hall.

Goldstein, H. (1986) Multilevel mixed linear model analysis using iterative generalised least squares. *Biometrika*, 73: 43-56.

Goldstein, H. (1995) *Multilevel statistical models*. 2nd ed. London: Edward Arnold.

Goldstein, H. (2003) *Multilevel statistical models.* Edward Arnold, London Wiley, New York.

Goldstein, H., and Rasbash, J. (1996) Improved approximations for multilevel models with binary responses. *Journal of the Royal Statistical Society A*, 159: 505-513.

Goldstein, H., and Woodhouse, G. (1996) Multilevel models with missing data. Eleventh International workshop on statistical modelling. Orvieto, Italy.

Goldstein, H., Browne, W., and Rasbash, J. (2002) Tutorial in biostatistics: multilevel modeling of medical data. *Statistics in Medicine*, 21: 3291-3315.

Gravelle, H. (1998) How much of the relation between population mortality and unequal distribution of income is a statistical artefact? *British Medical Journal*, 316: 382-385.

Haggett, P. (1965) *Locational Analysis in Human Geography*. Edward Arnold, London.

Health, Nutrition and Population data bank, (http://devdata.worldbank.org/hnpstats)

Hedeker, D., and Gibbons, R.D. (2005) *Applied longitudinal data analysis*. New York, John Wiley & Sons. Chapter 5: Mixed-effects polynomial regression models for continuous outcomes. P159-194.

Henderson, C., Liu, X., Diez Roux, AV., Link, B.G., and Hasin, D. (2004) The effects of US state income inequality and alcohol policies on symptoms of depression and alcohol dependence. *Social Science and Medicine*, 58: 565-575.

Hendryx, M.S., Ahern, M.M., Lovrich, N.P., and McCurdy, A.R. (2002) Access to health care and community social capital. *Health Service Research*, 31: 85-101.

Heylighen, F. (1992) A Cognitive-systemic reconstruction of Maslow's theory of self-actualization, *Behavioural Science*, 37: 39-58.

Human Development Report (2001) *Making new technologies work for human development.* New York: United Nations Development Programme.

Idler, E.L., and Benyamini, Y. (1997) Self-rated health and mortality: a review of twenty-seven community studies. *Journal of Health Social Behaviors*, 38: 21-37.

Inglehart, R. (1997) *Modernization and postmodernization: cultural, economic and political change in 43 societies.* Princeton University Press, Princeton.

Inglehart, R. *et al.*, (Principle investigators) (2000) *World values surveys and European values surveys, 1981–1984, 1990–1993, and 1995–1997.* Inter-university Consortium for Political and Social Research, Institute for Social Research, University of Michigan.

Inglehart, R., Basanez, M., Diez-Medrano, J., Halman, L., and Luijkx, R. (2004) *Human beliefs and values: a cross-cultural sourcebook based on the 1999-2002 values surveys.* Siglo XXI Editores. Maxico.

Johnston, R.J., Pattie, C.J., Dorling, D.F.L., MacAllister, I., Tunstall, H., and Rossiter D.J. (2001) Housing tenure, local context, scale and voting in England and Wales, 1997. *Electoral Studies*, 20: 195-216.

Jones, B., and Nagin, D.S. (2006) Advances in group-based trajectory modeling and a SAS procedure for estimating them. Submitted.

Jones, B.L., Nagin, D.S., and Roeder, K. (2001) A SAS procedure based on mixture models for estimating developmental trajectories. *Sociological Methods and Research*, 29: 374-393.

Jones, K., and Bullen, N. (1994) Contextual models of urban house prices: a comparison of fixed- and random-coefficient models developed by expansion. *Economic Geography*, 70: 252–272.

Jones, K., Duncan, C., and Twigg L. (2004) Evaluating the absolute and relative income hypothesis in an exploratory analysis of deaths in the Health and Lifestyle Survey. In: Boyle P., Curtis S., Graham E., eds. *The geography of health inequalities in the developed world.* London, United Kingdom: Ashgate Press.

Joung, I.M., van de Mheen, H., Stronks, K., van Poppel, F.W., and Mackenbach, J.P. (1994) Differences in self-reported morbidity by marital status and by living arrangement. *International Journal of Epidemiology*, 23: 91-97.

Judge, K., Mulligan, J.A., and Benzeval, M. (1998) Income inequality and population health. *Social Science and Medicine*, 46: 567-579.

Kaplan, G.A., Pamuk, E.R., Lynch, J.W., Cohen, R.D., Balfour, J.L., (1996) Inequality in income and mortality in the United States: analysis of mortality and potential pathways. *British Medicine Journal*, 312: 999-1003.

Kawachi, I., and Berkman, L. (2000) Social cohesion, social capital and health. In: Berkman, L.F., and Kawachi, I. eds. *Social Epidemiology*. New York, NY: Oxford University Press: P174-190.

Kawachi, I., and Kennedy, B.P. (1999) Income inequality and health: pathways and mechanisms. *Health service research*, 34: 215-217.

Kawachi, I., Kennedy, B.P., Lochner, K., and Prothrow-Stith D. (1997) Social capital, income inequality and mortality. *American Journal of public health*, 87: 1491-1498.

Kawachi, I., Kim, D., Coutts, A., and Subramanian, S.V. (2004) Commentary: Reconciling the three accounts of social capital. *International Journal of Epidemiology*, 33: 1-9.

Kennedy, B.P., Kawachi, I. Glass, R., and Prothrow-Stith, D. (1998) Income distribution, socioeconomic status and self-rated health in the United States. *British Medical Journal*, 317: 917-921.

Kennedy, B.P., Kawachi, I., and Brainerd, E. (1998) The role of social capital in the Russian mortality crisis. *World Development*, 26: 2029-2043.

Keys, A. (1950's) (http://www.epi.umn.edu/about/7countries/index.shtm)

Kubzansky, L.D., Sparrow, D., Vokonas, P. and Kawachi, I. (2001) Is the glass half empty or half full? A prospective study of optimism and coronary heart disease in the normative aging study. *Psychosomatic Medicine*, 63: 910-916.

Linden, J. van der, Drukker, M., Gunther, N., Feron, F., and Os, J. van. (2003) Children's mental health service use, neighbourhood socioeconomic deprivation, and social capital. *Social Psychiatry Psychiatric Epidemiology*, 38: 507-514.

Lindström, M., Merlo, J., and Ostergren, P-O. (2003)[a] Social capital and sense of insecurity in the neighbourhood: a population based multilevel analysis in Malmo, Sweden. *Social Science and Medicine*, 56: 1111–1120.

Lindström, M., Moghaddassi, M., and Merlo, J. (2003)[b] Social capital and leisure time physical activity: a population based multilevel analysis in Malmo, Sweden. *Journal of Epidemiology and Community Health*, 57: 23-28.

Lindström, M., Moghaddassi, M., Bolin, K., Lindgren, B., and Merlo, J. (2003)[c] Social participation, social capital and daily tobacco smoking: a population-based multilevel analysis in Malmö, Sweden. *Scandinavian Journal of Public Health*, 31: 444-450.

Lochner, K., Kawachi, I., Brennan R.T., and Buka S.L. (2002) Social capital and neighbourhood mortality rates in Chicago. *Social Science and Medicine*, 56: 1797-1805.

Lynch, J.W., Due, P., Muntaner, C., and Davey Smith, G. (2000) Social capital: is it a good investment strategy for public health. *Journal of Epidemiology and Community Health*, 54: 404-408.

Lynch, J.W., Davy-Smith, G., Hillemeier, M., Shaw, M., Raghunathan, T., and Kaplan, G. (2001) Income inequality, the psychosocial environment, and health: comparisons of wealthy nations. *Lancet*, 358: 194-200.

Lynch, J.W., Smith, G.D., Harper, S., Hillemeier, M., Ross, N., Kaplan, G.A., and Wolfson, M. (2004) Is income inequality a determinant of population health? Part 1: a systematic review. *The Milbank Quarterly*, 82: 5-99.

Macinko, J., Shi, L., Starfield, B., and Wulu, J. (2003) Income inequality and health: a critical review of the literature. *Medical Care Research and Review*, 60: 407-452.

Marmot, M.G., Smith, G.D., Stansfeld, S., Patel, C., North, F., Head, J., White, I., Brunner, E., Feeney, A. (1991) Health inequalities among British civil servants: the Whitehall II study. *Lancet*, 337: 1387-1393.

Meredith, W., and Tisak, J. (1990) Latent curve analysis. *Psychometrika*, 55: 107-122.

Merlo, J., Lynch, J.W., Yang, M., Lindstrom, M., Ostergren, P.O., Rasmusen, N.K., and Rastam, L. (2003) Effect of neighbourhood social participation on individual use of hormone replacement and anti-hypertensive medications: a multilevel analysis. *American Journal of Epidemiology*, 157: 774-783.

Moon, G, Gould, M., and colleagues (2000) Epidemiology: An introduction. Buckingham: Open University Press.

Muntaner, C. (1999) Social mechanisms, race and social epidemiology. *American Journal of Epidemiology*, 150: 121–126.

Muntaner, C., Lynch, J., and Oates, G.L. (1999) The social class determinants of income inequality and social cohesion. *International Journal of Health Services*, 29: 699-732.

Muth′en, B. (2003) Statistical and substantive checking in growth mixture modeling. *Psychological Methods*, 8: 369–377.

Muthén, B. (1989) Latent variable modelling in heterogeneous populations. *Psychometrika*, 54: 557–585.

Nagin, D.S. and Land, K.C. (1993) Age, criminal careers, and population heterogeneity: specification and estimation of nonparametric mixed Poisson model. *Criminology*, 31: 327-362.

Nagin, D., and Tremblay, R.E. (2001) Analyzing developmental trajectories of distinct but related behaviors: a group-based method. *Psychological Methods*, 6: 18-34.

Nagin, D.S. (2005) *Group-based modelling of development.* Harvard University Press.

Omran, A.R. (1971) The epidemiologic transition -- A theory of the epidemiology of population change. *The Milbank Memorial Fund Quarterly*, 49: 509-538.

Ostir, G.V., Markides, K.S., Peek, M.K., Goodwin, J.S. (2001) The association between emotional well-being and the incidence of stroke in older adults. *Psychosomatic Medicine*, 63: 210-215.

Pickles, A., and Angold, A. (2003) Natural categories or fundamental dimensions: on carving nature at the joints and the rearticulating of psychopathology. *Development and Psychopathology*, 15:529-551.

Popay, J. (2000) Social capital: the role of narrative and historical research. *Journal of Epidemiology and Community Health*, 54: 401.

Pritchett, L., and Summers, L.H. (1996) Wealthier is healthier. *Journal of Human Resources*, 31: 841-868.

Putnam, R.D. (1995) Bowling alone: America's declining social capital. *The Journal of Democracy*, 6: 65-78.

Rasbash, J., Steele, F., Browne, W., and Prosser, B. (2004) *A User's Guide to MLwiN.* Version 2.02, Centre for Multilevel Modeling, University of Bristol, UK.

Robinson, W. (1950) Ecological correlations and the behavior of individuals. *American Sociological Review*, 15: 351-357.

Rodgers, G. (1979) Income and Inequality as Determinants of Mortality: an international cross-section analysis. *Population Studies*, 33: 343-351.

Rodriguez, G., and N. Goldman (2001) Improved estimation procedures for multilevel models with binary response: a case-study, *Journal of the Royal Statistical Society A*, 164: 339-355.

Salomon, J.A., Murray, C. J. L. (2002) The epidemiologic transition revisited: compositional models for cause of death by age and sex. *Population and development review*, 28: 205-228.

Sampson, R. J., Raudenbush, S.W., and Earls, F. (1997) Neighbourhoods and violent crime: a multilevel study of collective efficacy. *Science*, 277: 918-924.

SAS Institute. Year 2004. *The SAS System 8.* Cary, NC: SAS Institute, Inc.

Schyns, P. (2002) Wealth of nations, individual income and life satisfaction in 42 countries: a multilevel approach. *Social Indicators Research*, 60: 5-40.

Shaw, M., Dorling, D. and Mitchell, R. (2002) *Health, place and society*, Pearson Education, Harlow.

Shy, C.M. (1997) The failure of academic epidemiology: witness for the prosecution. *American Journal of Epidemiology*, 145: 479-484.

Singer, J.D., and Willett, J.B. (2003) *Applied longitudinal data analysis. Modelling change and event occurrence.* Oxford University press.

Soobader, M.J., and LeClere, F.B. (1999) Aggregation and the measurement of income inequality: effects on mortality. *Social Science and Medicine*. 48: 733-744.

Spiegelhalter, D.J., Best, N. G, Garlin, B. P., and van der Linde, A. (2002) Bayesian measures of model complexity and fit (with discussion and rejoinder). *Journal of the*

Royal Statistical Society B, 64: 583-639.

Subramanian, S.V., Kawachi, I., and Kennedy, B.P. (2001) Does the state you live in make a difference? Multilevel analysis of self-rated health in the US. *Social Science and Medicine*, 53: 9-19.

Subramanian, S.V., Kim, D.J., and Kawachi, I. (2002) Social trust and self-rated health in US communities: a multilevel analysis. *Journal of Urban Health*, 79 (Supplement 1): S21-S34.

Subramanian, S.V., and Kawachi, I. (2003) The association between state income inequality and worse health is not confounded by race. *International Journal of Epidemiology*, 32: 1022-1028.

Subramanian, S.V., Blakely, T., and Kawachi, I. (2003) Income inequality as a public health concern: where do we stand? Commentary on "Is exposure to income inequality a public health concern?". *Health Service and Research*, 38: 153–167.

Subramanian, S.V., Degaldo, I., and Jadue, L. et al. (2003) Income inequality and health: multilevel analysis of Chilean communities. *Journal of Epidemiology and Community Health*, 57: 844-848.

Subramanian, S.V., and Kawachi, I. (2004) Income inequality and health: what have we learned so far? *Epidemiological Review*, 26: 78-91.

Sudman, S. (1976) *Applied sampling.* Academic Press, p.193.

Titterington, D.M., Smith, A.F., and Makov, U.E. (1985) Analysis of finite mixture distributions. New York: Wiley.

U.S. Department of Health and Human Services Centres for Disease Control (CDC) and Prevention National Centre for Health Statistics series 2, No 129. Method for constructing complete annual US life Tables.

Veenhoven, R. (1989) *How harmful is happiness? Consequences of enjoying life or not.* Universitaire Pers Rotterdam, 1989, The Netherlands, Chapter 6 Does happiness bind? Marriage chances of the unhappy. P44-60.

Veenhoven, R. (1991) "Is happiness relative?" *Social Indicators Research*, 24: 1-34.

Veenstra, G. (2005) Location, location, location: contextual and compositional health effects of social capital in British Columbia, Canada. *Social Science and Medicine*, 60: 2059-2071.

Veenstra, G., Luginaah, I., Wakefield, S., Birch, S., Eyles, J., and Elliott, S. (2005) Who you know, where you live: social capital, neighbourhood and health. *Social Science and Medicine*, 60: 2799-2818.

Wagstaff, A., and van Doorslaer E. (2000) Income inequality and health: what does the literature tell us? *Annual review of public health*, 21: 543-567.

Wasserman, L. (2000) Bayesian model selection and model averaging. *Journal of*

Mathematical Psychology, 44: 92-107.

Wen, M., Browning, C.R., and Cagney, K.A. (2003) Poverty, affluence, and income inequality: neighbourhood economic structure and its implications for health. *Social Science and Medicine*, 57: 843-860.

Wen, M., and Christakis, N.A. (2005) Neighbourhood effects on post hospitalization mortality: a population-based cohort study of the elderly in Chicago. *Health Services Research*, 40: 1108-1127.

Wilcox, V.L., Kasl, S.V., and Idler, E.L. (1996) Self-rated health and physical disability in elderly survivors of a major medical event. *Journal of Gerontology B Psychological Science and Social Science*, 51: 96-104.

Wilkinson, R.G (1996) *Unhealthy societies: the afflictions of inequality.* London Routledge.

Wilkinson, R.G (2001) *Mind the gap: hierarchies, health and human evolution.* Yale University Press.

Willett, J. B., & Sayer, A. G. (1994). Using covariance structure analysis to detect correlates and predictors of individual change over time. *Psychological Bulletin*, 116: 363–381.

World Health Organization, (2001) Men, ageing and health, achieving health across the life span (http://whqlibdoc.who.int/hq/2001/WHO_NMH_NPH_01.2.pdf).

Xi, G., McDowell, I., Nair, R., and Spasoff, R. (2005) Income inequality and health in Ontario: a multilevel analysis. *Canadian Journal of Public Health. Revue Canadienne de Sante Publique*, 96: 206-211.

Appendix 1 Data to be used from the World values survey

Structural variables

- Country code

Code	Country	Code	Country	Code	Country
1	France	34	E Germany	68	Dominic Rep
2	Britain	35	Slovenia	69	Bangladesh
3	W Germany	36	Bulgaria	70	Indonesia
4	Italy	37	Romania	71	
5	Netherlands	38	Pakistan	72	Albania
6	Denmark	39	China 90	73	Colombia
7	Belgium	40	Taiwan	74	
8	Spain	41	Portugal	75	Basque
9	Ireland	42	Austria	76	
10	N Ireland	43	Greece	77	
11	USA	44	Turkey	78	Andalusia
12	Canada	45	Moscow	79	Galicia
13	Japan	46	Lithuania	80	Valencia
14	Mexico	47	Latvia	81	Serbia
15	S Africa	48	Estonia	82	Montenegro
16	Hungary	49	Ukraine	83	Macedonia
17	Australia	50	Russia	84	Croatia
18	Norway	51	Peru	85	Slovakia
19	Sweden	52	EI Salvador	86	
20	Tambov	53	Venezuela	87	
21	Iceland	54	Uruguay	88	New Zealand
22	Argentina	55		89	Egypt
23	Finland	56	Ghana	90	Morocco
24	S Korea	57		91	Iran
25	Poland	58	Philippines	92	Jordan
26	Switzerland	59	Israel	93	Bosnia
27	Puerto Rico	60		94	
28	Brazil	61	Moldova	95	
29	Nigeria	62	Georgia	96	
30	Chile	63	Armenia	97	
31	Belarus	64	Azerbaijan	98	
32	India	65		99	
33	Czech	66			
		67			

- Survey year

Survey number	Year
1	1981-82 wave
2	1990-91 wave
3	1995-1997 wave
4	1999-2001 wave

- Interview number

A 4-digit number identifying each respondent in the given country. Most of the variables in the questionnaire are 1-digit, using "9" as the missing data code; if the variable has 8 or more categories, "99" is the missing data code.

- Weight

In each country, the investigators were asked to provide a 4-digit weight variable to correct their sample to reflect national distributions of key variables. If no weighting was necessary, each case was simply codes as "1.00."

Dependent variables

Number	Question	Recoding
V11	All in all, how would you describe your state of health these days? Would you say it is	1. Very good 2. Good 3. Fair 4. Poor 5. Very poor 9. Don't know
V10	Taking all things together, would you say you are:	1. Very happy 2. Quite happy 3. Not very happy 4. Not at all happy 9. Don't know
V65	All things considered, how satisfied are you with your life as a whole these days? Please use this card to help with your answer.	1. Dissatisfied 2. 3. 4. 5. 6. 7. 8. 9. 10. Satisfied 99. Don't know

Independent variables

Number	Question	Recoding
V216	This means you are ___ years old.	1. 18-24 years 2. 25-34 years 3. 35-44 years 4. 45-54 years 5. 55-64 years
V215	Can you tell me your year of birth, please? 19??_	
V89	Are you currently	1. Married 2. Living together as married 3. Divorced 4. Separated 5. Widowed 6. Single
V227	Here is a scale of income. We would like to know in what group your household is, counting all wages, salaries, pensions and other incomes that come in. Just give the letter of the group your household falls into, before taxes and other deductions.	1. C (Lowest Decile) 2. D 3. E 4. F 5. G 6. H 7. I 8. J 9. K 10. L (Highest Decile) 98. No answer.